

1/5/1 (Item 1 from file: 351)
DIALOG(R)File 351:Derwent WPI
(c) 2004 Thomson Derwent. All rts. reserv.

012857098 **Image available**
WPI Acc No: 2000-028931/ 200003
XRPX Acc No: N00-021955

Memory conservation method of computer network e.g. internet proxy, web browsing server - involves generating low resolution version of object which occupies less memory space than complete resolution version and storing it

Patent Assignee: IBM CORP (IBM) ; INT BUSINESS MACHINES CORP (IBM)
Inventor: LAMAI R O; ROBINSON J T
Number of Countries: 002 Number of Patents: 004
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 11296426	A	19991029	JP 9926262	A	19990203	200003 B
US 6343350	B1	20020129	US 9823793	A	19980213	200210
US 6378053	B1	20020423	US 9823793	A	19980213	200232
			US 2000715985	A	20001120	
JP 3504526	B2	20040308	JP 9926262	A	19990203	200418

Priority Applications (No Type Date): US 9823793 A 19980213; US 2000715985
A 20001120

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 11296426	A		27	G06F-012/00	
US 6343350	B1			G06F-012/00	
US 6378053	B1			G06F-012/00	Div ex application US 9823793
JP 3504526	B2		17	G06F-012/00	Previous Publ. patent JP 11296426

Abstract (Basic): JP 11296426 A

NOVELTY - The low resolution version of an object which occupies less energy space than complete resolution version of object is generated and stored in the memory of internet proxy (160). DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for the object management method.

USE - In computer network e.g. internet proxy, web browsing server.

ADVANTAGE - Since low resolution version of object occupies less memory space, vital resource like memory space is efficiently used, thus economizing the memory quantity. DESCRIPTION OF DRAWING(S) - The figure shows the diagram of computer network. (160) Internet proxy.

Dwg.1/12

Title Terms: MEMORY; CONSERVE; METHOD; COMPUTER; NETWORK; WEB; SERVE;
GENERATE; LOW; RESOLUTION; VERSION; OBJECT; OCCUPY; LESS; MEMORY; SPACE;
COMPLETE; RESOLUTION; VERSION; STORAGE

Derwent Class: T01

International Patent Class (Main): G06F-012/00

International Patent Class (Additional): G06F-013/00

File Segment: EPI

1/5/2 (Item 1 from file: 347)
DIALOG(R)File 347:JAPIO
(c) 2004 JPO & JAPIO. All rts. reserv.

06354819 **Image available**
SAVING METHOD FOR STORAGE AREA SPACE, MANAGING METHOD FOR OBJECT AND PRODUCT

PUB. NO.: 11-296426 A]
PUBLISHED: October 29, 1999 (19991029)
INVENTOR(s): RICHARD OLVIRI LAMAILE
ROBINSON JOHN T
APPLICANT(s): INTERNATL BUSINESS MACH CORP <IBM>;
APPL. NO.: 11-026262 [JP 9926262]

FILED: February 03, 1999 (19990203)
PRIORITY: 23793 [US 23793], US (United States of America), February 13,
1998 (19980213)
INTL CLASS: G06F-012/00; G06F-013/00

ABSTRACT

PROBLEM TO BE SOLVED: To provide a method for saving a memory space by installing a first resolution edition of an object in a memory, generating a low resolution edition of the object using the memory space, which is less than the first resolution edition, preserving the low resolution edition in the memory and abolishing the first resolution edition from the memory.

SOLUTION: A client computer 100 is connected to a proxy 160 and the proxy 160 is connected to a web server 170. A data object is retrieved through a communication network and it is stored in the proxy 160 or a web browser and the like. When the data object is retrieved for the first time, the low resolution edition of the object is calculated and it is preserved with a high resolution edition. In a method for saving the memory space, the first resolution edition of the object is given to the memory and the low resolution edition of the object using the memory space, which is less than the first resolution edition, is generated, the low resolution edition is preserved in the memory and the first resolution edition is abolished from the memory.

COPYRIGHT: (C)1999, JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-296426

(43) 公開日 平成11年(1999)10月29日

(51) Int.Cl.⁶

G 0 6 F 12/00
13/00

識別記号

5 4 6
3 5 4

F I

G 0 6 F 12/00
13/00

5 4 6 K
3 5 4 D

審査請求 有 請求項の数47 O L (全 27 頁)

(21) 出願番号 特願平11-26262

(22) 出願日 平成11年(1999) 2 月 3 日

(31) 優先権主張番号 0 9 / 0 2 3 7 9 3

(32) 優先日 1998年 2 月13日

(33) 優先権主張国 米国 (U S)

(71) 出願人 390009531

インターナショナル・ビジネス・マシー
ズ・コーポレーション
INTERNATIONAL BUSIN
ESS MACHINES CORPO
RATION

アメリカ合衆国10504、ニューヨーク州
アーモンク (番地なし)

(72) 発明者 リチャード・オルヴィリ・ラマイレ
アメリカ合衆国10598 ニューヨーク州、
ヨークタウン・ハイツ、 カリフォルニ
ア・ロード 138

(74) 代理人 弁理士 坂口 博 (外 1 名)

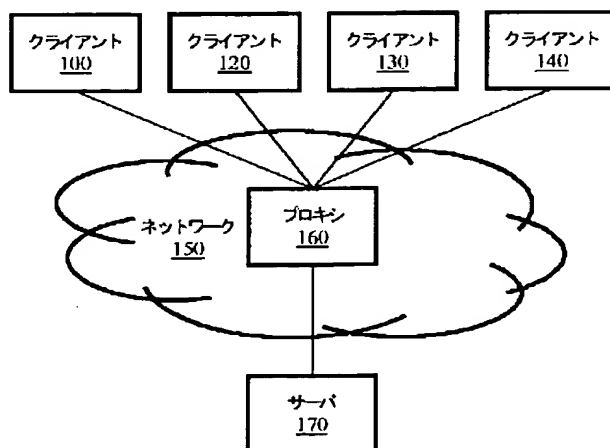
最終頁に続く

(54) 【発明の名称】 記憶域空間の節約方法、オブジェクトの管理方法及び製造物

(57) 【要約】

【課題】 メモリ空間を節約するための方法および製造物を提供する。

【解決手段】 システム（例えば、インターネット・プロキシ又はウェブ・ブラウザ）内のメモリ空間（キャッシュ形式であることが多い）は、新しいオブジェクトのための空間を作成するため完全解像度版が除去されるとき（例えば、存続時間の長さにより、使用近時性最小（LRU）置換方針を使用して）、データ・オブジェクトの低解像度版を保存することによって節約される。これは、メモリ空間の効率的な使用を提供する。多くの場合、データ・オブジェクトの低解像度版は十分な使用品質を有する。さらに、そのような場合でなくても、低解像度版を素早く入手し、しばらくの遅延の後に、ユーザ又はアプリケーションの要求によって完全解像度版を入手することが有用な場合がある。本発明を実現した方法、装置、及び製造物の特定の有利な実施例が提示される。



1

【特許請求の範囲】

【請求項 1】複数のオブジェクトをメモリに記憶するユニットで記憶域空間の量を節約する方法であって、オブジェクトの 1 つの第 1 の解像度版をメモリに有するステップと、

前記オブジェクトの 1 つの少なくとも 1 つの低解像度版であって、前記第 1 の解像度版よりも少ないメモリ空間を使用する解像度版を発生するステップと、

前記低解像度版をメモリに保存するステップと、

前記第 1 の解像度版をメモリから廃棄するステップとを含む、

記憶域空間の節約方法。

【請求項 2】複数のオブジェクトをキャッシュ内で管理する方法であって、

存在する非多解像度オブジェクト及び少なくとも 1 つの多解像度オブジェクトのためにキャッシュ・ディレクトリ・エントリーを有するキャッシュ・ディレクトリを作成し、前記多解像度オブジェクトの各キャッシュ・ディレクトリ・エントリーがそのオブジェクトのすべての存在する解像度版を指すポインタを有するようにするステップと、

前記エントリーの順序リストを形成するステップとを含む、

オブジェクトの管理方法。

【請求項 3】前記形成するステップが、特定のオブジェクトが使用された最後の使用時点に基づく、請求項 2 に記載のオブジェクトの管理方法。

【請求項 4】前記順序リストがオブジェクト使用基準の頻度に従って形成される、請求項 2 に記載のオブジェクトの管理方法。

【請求項 5】各オブジェクトを現在の解像度版で記憶するステップと、

前記順序リスト内の最低優先順位オブジェクト y に対応するキャッシュ・ディレクトリ・エントリーが、多解像度オブジェクトであるかどうかを確認するステップと、オブジェクト y が多解像度オブジェクトであれば、オブジェクト y の低減解像度版を発生及び記憶するステップと、

前記キャッシュ中に自由空間を作成するためにオブジェクト y の前記現在の解像度版を除去するステップと、オブジェクト y の前記キャッシュ・ディレクトリ・エントリーを最高優先順位オブジェクトに設定するステップと、

前記自由空間の量がオブジェクト x を記憶するのに十分であるかどうかを決定するステップと、

前記自由空間の量がオブジェクト x を記憶するのに十分でなければ、

前記自由空間の量がオブジェクト x を記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リスト中の次の最低優先順位オブジェクトのた

2

めに確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、

前記自由空間にオブジェクト x を記憶するステップとを含む、

請求項 2 に記載のオブジェクトの管理方法。

【請求項 6】 y が多解像度オブジェクトでなければ、

前記自由空間の量を増加するためにオブジェクト y を前記キャッシュから除去するステップと、

前記自由空間の量がオブジェクト x を記憶するのに十分であるかどうかを決定するステップと、

前記自由空間の量がオブジェクト x を記憶するのに十分でなければ、前記自由空間の量がオブジェクト x を記憶

するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リスト内の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、

オブジェクト x を前記自由空間に記憶するステップとを含む、

請求項 5 に記載のオブジェクトの管理方法。

20 【請求項 7】各オブジェクトを現在の解像度版で記憶するステップと、

前記順序リストの最後に置かれた使用近時性最小オブジェクトに対応する第 1 のキャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを確認するステップと、

オブジェクト y が多解像度オブジェクトであれば、

オブジェクト y の低減解像度版を発生及び記憶するステップと、

30 自由空間を作成するように、オブジェクト y の現在の高解像度版を除去するステップと、

オブジェクト y の前記第 1 のキャッシュ・ディレクトリ・エントリーを前記順序リストの始めに配置することによって前記第 1 のキャッシュ・ディレクトリ・エントリーを使用近時性最大オブジェクトに設定するステップと、

前記自由空間の量がオブジェクト x を記憶するのに十分であるかどうかを決定するステップと、

もし十分でなければ、前記自由空間の量がオブジェクト x を記憶するのに十分となるまで、キャッシュ・ディレ

40 クトリ・エントリーの前記順序リストにおける次の使用近時性最小オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクト x を前記自由空間に記憶するステップとを含む、

請求項 2 に記載のオブジェクトの管理方法。

【請求項 8】 y が多解像度オブジェクトでなければ、

前記自由空間の量を増加するためオブジェクト y を前記キャッシュから除去するステップと、

前記自由空間がオブジェクト x を記憶するのに十分であるかどうかを決定するステップと、

3

前記自由空間の量が十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまでキャッシュ・ディレクトリ・エントリーの前記順序リストにおける次の使用近時性最小オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、

オブジェクトxを前記自由空間に記憶するステップとを含む、

請求項7に記載のオブジェクトの管理方法。

【請求項9】前記管理するステップが、前記多解像度オブジェクトの少なくとも1つのオブジェクト解像度を低減するステップを含む、請求項2に記載のオブジェクトの管理方法。

【請求項10】オブジェクト・キャッシュ内のオブジェクトを管理する方法であって、

現在前記キャッシュ中に存在しない新しいオブジェクトを記憶するために空間を作成するステップと、

多解像度オブジェクトの位置を決定することによって前記キャッシュ中の特定のオブジェクトを置換のために選択するステップと、

前記多解像度オブジェクトの最高解像度版を廃棄するステップと、

前記オブジェクトの次の最低解像度版があれば、それを保持するステップと、

前記多解像度オブジェクトの次の最低解像度版の優先順位を増加するステップと、

前記多解像度オブジェクトが後の置換に対して即時に適格となることを防止するステップとを含む、

オブジェクトの管理方法。

【請求項11】増加する前記ステップが、特定のオブジェクトが使用された最後の使用時点に基づいて順序リストを形成するステップを含む、請求項10に記載のオブジェクトの管理方法。

【請求項12】複数のオブジェクトをキャッシュに記憶するシステムでオブジェクトが多解像度タイプであるかどうかを決定する方法であって、

前記システム中で前記現在理解された多解像度タイプ及び各オブジェクト・タイプのために低減解像度版を発生する前記プロシージャをリストするテーブルを作成するステップと、

オブジェクトの前記タイプと前記テーブル中にリストされた複数のタイプの1つとの間で一致を探索するステップと、

一致が発見されなければ、前記オブジェクトが多解像度オブジェクトではないと表示するステップと、

一致が発見されれば、前記オブジェクトが多解像度オブジェクトであること、及び前記テーブル中にリストされた解像度低減プロシージャが前記オブジェクトの低解像度版を発生するために使用できることを表示するステップとを含む、

4

多解像度オブジェクトの決定方法。

【請求項13】第1のコンピュータ・システムが、所与の名前を有する多解像度オブジェクトを所与の解像度レベルで渡すように第2のコンピュータ・システムへ要求する方法であって、

前記第1のコンピュータ・システムから前記第2のコンピュータ・システムへ、前記多解像度オブジェクトを要求するメッセージを送るステップと、

前記第2のコンピュータ・システムが、前記要求されたオブジェクトを、要求された解像度レベルで有するとき、前記要求されたオブジェクトを前記第1のコンピュータ・システムへ渡すステップと、

前記第2のコンピュータ・システムが、前記要求されたオブジェクトを、高解像度レベルで有し、また前記要求された解像度レベルを発生するのに必要な前記プロシージャを有するとき、前記要求されたオブジェクトを前記第1のコンピュータ・システムへ渡すステップと、

前記第2のコンピュータ・システムが、前記要求されたオブジェクトを、前記要求された解像度レベルで有しないで、前記オブジェクトが存在するとき、前記要求された解像度を発生する方法を有しないとき、前記要求されたオブジェクトが利用できないことを示すメッセージを前記第1のコンピュータ・システムへ送るステップとを含む、

オブジェクトの要求方法。

【請求項14】現在前記キャッシュ中に存在しない新しいオブジェクトxを記憶する自由空間を作成するために、オブジェクト・キャッシュ内のオブジェクトを管理する方法であって、

各多解像度オブジェクトを、低解像度版によって占められるキャッシュ空間の量が高解像度版によって占められる前記空間の量よりも小さくなるような複数の解像度版で記憶するステップと、

存在する非多解像度オブジェクト及び前記少なくとも1つの多解像度オブジェクトのためにエントリーを有するキャッシュ・ディレクトリを作成し、各キャッシュ・ディレクトリ・エントリーが前記少なくとも1つの多解像度オブジェクトのすべての存在する解像度版を指すポインタを有するようにするステップと、

特定のオブジェクトが使用された最後の使用時点に基づく前記順序基準に従って前記エントリーの順序リストを形成ステップと、

前記順序リストの前記終わりにある使用近時性最小オブジェクトyに対応する前記キャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、

yが多解像度オブジェクトでないとき、オブジェクトyを除去して、新しく増加した自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを確認するステップと、

前記順序リストの前記終わりにある使用近時性最小オブジェクトyに対応する前記キャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、

yが多解像度オブジェクトでないとき、オブジェクトyを除去して、新しく増加した自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを確認するステップと、

新しく増加した前記自由空間の量がオブジェクトxを記憶するのに十分でないとき、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リスト内の次の使用近時性最小オブジェクトのために決定及び確認の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、

yが多解像度オブジェクトであるとき、前記オブジェクトの複数の解像度版が現在記憶されているかどうかを決定するステップと、

1つの解像度版のみが現在記憶されているとき、オブジェクトyの低減解像度版を発生及び記憶するステップと、

オブジェクトyの前記現在の高解像度版を除去するステップと、

前記オブジェクトyのキャッシュ・ディレクトリ・エントリを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリを使用近時性最大オブジェクトに設定し、前記新しく増加した自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを確認するステップと、

新しく増加した前記自由空間の量がオブジェクトxを記憶するのに十分でないとき、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リストにある前記次の使用近時性最小オブジェクトのために決定及び確認の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、

複数の解像度版が現在記憶されていれば、オブジェクトyの前記現在の高解像度版を除去するステップと、

オブジェクトyの前記キャッシュ・ディレクトリ・エントリを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリを前記使用近時性最大オブジェクトに設定し、前記新しく増加した自由空間がオブジェクトxを記憶するのに十分であるかどうかを確認するステップと、

十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リストにあるオブジェクトyの前記次の最高解像度版のために除去、設定、及び確認の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、

オブジェクトの管理方法。

【請求項 1 5】複数のオブジェクトをキャッシュに記憶するシステムでオブジェクトが多解像度タイプであるかどうかを決定する方法であって、

前記現在理解されている多解像度タイプをリストするテーブルを前記システム中に作成するステップと、

オブジェクトの前記タイプと前記テーブルにリストされた前記タイプとの間の一致を探索するステップと、

一致が発見されないとき、

前記オブジェクトが多解像度オブジェクトではないことを表示するステップと、

一致が発見されたとき、前記オブジェクトが多解像度オブジェクトであることを表示するステップとを含む、オブジェクトの決定方法。

【請求項 1 6】前記テーブルが、各オブジェクト・タイプのために低減解像度版を発生するプロシーダをリストしているとき、

10 前記一致が発見されると、前記テーブル内の情報から、前記オブジェクトの低減解像度版を発生するために特定の解像度低減プロシーダが使用されるべきであることを表示するステップを含む、

請求項 1 5 に記載のオブジェクトの決定方法。

【請求項 1 7】ネットワーク・オブジェクト・キャッシュ内のオブジェクトを変更する方法であって、現在前記キャッシュ中に存在しないオブジェクト'x'によって必要とされる記憶域の前記キャッシュ・サイズ'Sx'バイトを決定するステップと、

20 前記キャッシュ内の現在の自由空間の量'S_FREE'が'Sx'よりも大きい、それに等しいかどうかを決定するステップと、

'S_FREE'が'Sx'よりも大きい場合、さらにキャッシュ・メモリの'Sx'バイトを'x'へ割り振り、'x'を前記割り振られたキャッシュに記憶するステップと、

'x'のためにキャッシュ・ディレクトリ・エントリ(CDE)を作成し、前記CDEをハッシュ・チェーンに配列するステップと、

30 前記CDEを現在の使用近時性最大オブジェクトに設定するステップと、

'S_FREE'を'S_FREE'-'Sx'に設定するステップとを含む、

オブジェクトの変更方法。

【請求項 1 8】'Sx'が'S_FREE'よりも大きいとき、前の使用近時性最小オブジェクト'y'を発見するステップと、

オブジェクト'y'が多解像度オブジェクトであるかどうかを確認するステップと、もしオブジェクト'y'が多解像度オブジェクトでなければ、

40 'y'のCDEを除去して'Sy'バイトから構成されるオブジェクト'y'へ割り振られたキャッシュ空間の量を解放し、'S_FREE_NEW'='S_FREE'+ 'Sy'を形成するステップと、

'S_FREE_NEW'が'Sx'よりも大きいかどうかを決定し、もし大きければ、'x'を記憶し、もし大きくなければ、次の使用近時性最小オブジェクト'y'のために前記確認ステップを反復するステップとを含む、

請求項 1 7 に記載のオブジェクトの変更方法。

【請求項 1 9】オブジェクト'y'が多解像度オブジェクトであるとき、

オブジェクト'y'が前記キャッシュ内の解像度'レベル0'版で存在するかどうかを決定し、もしそれが存在すれば、前記解像度'レベル0'位置フィールドをヌルにして、'S_FREE_NEW'='S_FREE'+ 'Sy'を形成するステップと、

'S_FREE_NEW'が'Sx'よりも大きいかどうかを決定し、もし大きければ、'x'を記憶し、もし大きくなければ、次の使用近時性最小オブジェクト'y'のために前記確認ステップを反復するステップとを含む、

請求項 1 8 に記載のオブジェクトの変更方法。

【請求項 2 0】オブジェクト'y'が特定の多解像度オブジェクトであるとき、オブジェクト'y'の解像度'レベル1'版が前記キャッシュに存在するかどうかを決定し、もしそれが存在すれば、前記解像度'レベル0'版のためにレベル1空間を解放し、前記解像度'レベル0'版のためにオブジェクト'y'の前記CDEをヌルにし、'S_FREE'を'S_FREE'_NEW2='S_FREE'+ 'Sy'_SIZE_0へ増加し、オブジェクト'y'のCDEを使用近時性最大オブジェクトに設定し、'S_FREE_NEW2'が'Sx'よりも大きいかどうかを決定し、もし大きければ、'x'を記憶し、もし大きくなければ、次の使用近時性最小オブジェクト'y'のために前記確認するステップを反復するステップを含む、

請求項 1 8 に記載のオブジェクトの変更方法。

【請求項 2 1】解像度1版が発見されなかったとき、オブジェクト'y'に関連づけられた'タイプ'に関連づけられた解像度低減プロシージャを、オブジェクト'y'の前記解像度'レベル0'版へ適用して、オブジェクト'y'の解像度'レベル1'版を発生するステップと、前記キャッシュ内の'y'_RES_0に等しい空間の量を解放し、前記解像度'レベル0'位置をヌルにするステップと、オブジェクト'y'を解像度レベル版'y'_RES_1で前記キャッシュに記憶するステップと、前記CDE内の解像度1位置フィールドが前記レベル0位置を指すように設定するステップと、'S_FREE'を'S_FREE'+ 'Sy'_SIZE_0へ増加するステップと、'S_FREE'を'S_FREE'- 'Sy'_SIZE_1に減少するステップと、

オブジェクト'y'の'y'_CDEを前記使用近時性最大に設定するステップとを含む、

請求項 2 0 に記載のオブジェクトの変更方法。

【請求項 2 2】コンピュータ使用可能媒体を含み、複数のオブジェクトをキャッシュに記憶するシステムでキャッシュ空間の量を節約するコンピュータ読み取り可能プログラム・コード手段が前記コンピュータ使用可能媒体に実現された製造物であって、前記コンピュータ読み取り可能プログラム・コード手段は、

少なくとも1つの前記オブジェクトの少なくとも1つの

小サイズ低解像度版を発生するステップと、前記低解像度版を前記キャッシュに保存するステップと、前記少なくとも1つの前記オブジェクトの大サイズ高解像度版を廃棄するステップとをコンピュータに実行させる、コンピュータ使用可能媒体を含む製造物。

【請求項 2 3】前記オブジェクトの各々が第1解像度のデータを有するとき、前記オブジェクトの1つの前記第1解像度のデータが第1解像度削除方針に従っていつ削除されるかを決定するステップを前記コンピュータに実行させる、請求項 2 2 に記載のコンピュータ使用可能媒体を含む製造物。

【請求項 2 4】前記オブジェクトの各々が低解像度のデータを有するとき、前記オブジェクトの1つの前記低解像度のデータが低解像度削除方針に従っていつ削除されるかを決定するアクションを前記コンピュータに実行させる、請求項 2 2 に記載のコンピュータ使用可能媒体を含む製造物。

【請求項 2 5】複数のオブジェクトをキャッシュに記憶するシステムで前記有効なキャッシュ空間を節約する方法のステップを実行するために前記機械によって実行可能なプログラム命令を有形的に実現した、機械読み取り可能なプログラム記憶装置であって、前記方法のステップは、

少なくとも1つの前記オブジェクトの少なくとも1つの小サイズ低解像度版を発生するステップと、前記低解像度版をキャッシュに保存するステップと、前記少なくとも1つの前記オブジェクトの大サイズ高解像度版を廃棄するステップとを含む、プログラム記憶装置。

【請求項 2 6】前記オブジェクトの各々が第1のキャッシュ量を占める第1解像度のデータを有するとき、前記方法が、キャッシュの自由空間の量を増加して前記キャッシュに存在しない他のオブジェクトのための余地を作るため、前記オブジェクトの1つの前記第1解像度のデータが第1解像度削除方針に従っていつ削除されるかを決定するステップを含む、

請求項 2 5 に記載のプログラム記憶装置。

【請求項 2 7】前記オブジェクトの各々が小サイズ低解像度のデータを有するとき、前記方法が、オブジェクトの1つの前記小サイズ低解像度のデータが低解像度削除方針に従っていつ削除されるかを決定するステップを含む、請求項 2 5 に記載のプログラム記憶装置。

【請求項 2 8】各々の多解像度オブジェクトが最初に複数の解像度版で記憶されるとき、現在前記キャッシュに存在しない新しいオブジェクトxを記憶する空間を作成するために使用され、

前記順序リストの前記終わりにある最低優先順位オブジ

ェクトに対応する特定のキャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、

yが多解像度オブジェクトであるとき、オブジェクトyの複数の解像度版が現在記憶されているかどうかを確認するステップと、

オブジェクトyの複数の解像度版が現在記憶されているとき、

自由空間を形成するため前記オブジェクトyの現在の最高解像度版を除去するステップと、

オブジェクトyの前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって、オブジェクトyの前記キャッシュ・ディレクトリ・エントリーを最高優先順位オブジェクトに設定するステップと、

前記自由空間の量が前記オブジェクトxを記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、順序リスト内の次の最低優先順位オブジェクトのために決定、確認、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、

請求項2に記載のオブジェクトの管理方法。

【請求項29】オブジェクトyの1つの解像度版のみが現在記憶されているとき、

オブジェクトyの低減解像度版を発生及び記憶するステップと、

前記自由空間の量を増加するためオブジェクトyの前記現在の高解像度版を除去するステップと、

オブジェクトyの前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリーを前記最高優先順位オブジェクトに設定するステップと、

前記新しく増加された自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを計算するステップと、

もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リストにおける前記次の最低優先順位オブジェクトのために決定、確認、発生、除去、設定、配置、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、yが多解像度オブジェクトでないとき、

前記オブジェクトyを除去し、前記新しく増加された自由空間の量が前記オブジェクトxを記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記順序リスト内の前記次の最低優先順位使用オブジェクトのために決定、確認、発生、除去、及び計算の前記ステップを反復するステップとを含む、

請求項28に記載のオブジェクトの管理方法。

【請求項30】複数のオブジェクトをキャッシュに記憶するシステムでキャッシュ空間の量を節約する方法であって、

少なくとも1つの前記オブジェクトの少なくとも1つの小メモリ・サイズ版であってキャッシュ空間の第1の量を占める版を発生するステップと、

前記小メモリ・サイズ版をキャッシュに保存するステップと、

10 前記少なくとも1つの前記オブジェクトの高メモリ・サイズ版であってキャッシュ空間の前記第1の量よりも大きいキャッシュ空間の第2の量を占める版を廃棄するステップとを含む、

キャッシュ空間の節約方法。

【請求項31】キャッシュ内の複数のオブジェクトを管理する方法であって、

存在する非多解像度オブジェクト及び少なくとも1つの多解像度オブジェクトのためにエントリーを有するキャッシュ・ディレクトリを作成し、各キャッシュ・ディレクトリ・エントリーが前記少なくとも1つの多解像度オブジェクトのすべての存在する解像度版を指すポインタを有するようにするステップと、

前記エントリーの順序リストを形成するステップとを含む、

オブジェクトの管理方法。

【請求項32】前記順序リストが配列基準に従って形成される、請求項31に記載のオブジェクトの管理方法。

【請求項33】前記配列基準が、特定のオブジェクトが使用された最後の使用時点に基づく、請求項32に記載のオブジェクトの管理方法。

30 【請求項34】各々の多解像度オブジェクトが最初に複数の解像度版で記憶され、各々の解像度版が異なった量のキャッシュ空間を使用しているとき、現在前記キャッシュに記憶されていない新しいオブジェクトxを記憶する空間を作成するために使用され、

前記順序リストの前記終わりにある使用近時性最小オブジェクトyに対応する特定のキャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、

yが多解像度オブジェクトであるとき、前記オブジェクトの複数の解像度版が現在記憶されているかどうかを確認するステップと、

複数の解像度版が現在記憶されているとき、前記オブジェクトyの前記現在の最高解像度版を除去するステップと、

オブジェクトyの前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリーを前記使用近時性最大オブジェクトに設定するステップと、

50 新しく増加した自由空間の量が前記オブジェクトxを記憶するのに十分であるかどうかを計算し、もし十分でな

ければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、前記順序リスト内の前記次の使用近時性最小オブジェクトのために決定、確認、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、請求項33に記載のオブジェクトの管理方法。

【請求項35】1つの解像度版のみが現在記憶されているとき、

オブジェクトyの低減解像度版を発生及び記憶するステップと、

前記オブジェクトyの前記現在の高解像度版を除去するステップと、

前記順序リストの前記始めにオブジェクトyの前記キャッシュ・ディレクトリ・エントリを配置することによって前記キャッシュ・ディレクトリ・エントリを前記使用近時性最大オブジェクトに設定するステップと、前記新しく増加した自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを計算するステップと、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リストにある前記次の使用近時性最小オブジェクトのために決定、確認、発生、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、もしyが多解像度版オブジェクトでなければ、前記オブジェクトyを除去し、前記新しく増加した自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、前記順序リスト内の前記次の使用近時性最小オブジェクトのために決定、確認、発生、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、請求項34に記載のオブジェクトの管理方法。

【請求項36】キャッシュ及びキャッシュ・マネージャを備えた装置であつて、前記キャッシュ・マネージャは、複数のオブジェクトが前記キャッシュに記憶される際のキャッシュ空間の量を節約するキャッシュ空間節約手段を含む、装置。

【請求項37】CPU、ネットワーク・オブジェクト縮減手段、及びメモリ・ディスクを備えた、請求項36に記載の装置。

【請求項38】前記キャッシュがキャッシュ・ディレクトリを備えた、請求項36に記載の装置。

【請求項39】オブジェクトの各々へ割り当てられた優先順位に従って前記オブジェクトの順序リストを形成する前記ステップを含む、請求項1に記載の記憶域空間の節約方法。

【請求項40】割り当てられた前記優先順位がオブジェ

クト使用の頻度に基づく、請求項39に記載の記憶域空間の節約方法。

【請求項41】各オブジェクトを現在の解像度版で記憶するステップと、

最低優先順位オブジェクトyが多解像度オブジェクトであるかどうかを確認するステップと、

オブジェクトyが多解像度オブジェクトであるとき、

オブジェクトyの低減解像度版を発生及び記憶するステップと、

10 自由空間の量を前記メモリに作成するためオブジェクトyの前記現在の解像度版を除去するステップと、

オブジェクトyの優先順位を最高優先順位に設定するステップと、

前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、

前記自由空間の量がオブジェクトxを記憶するのに十分でないとき、前記自由空間の量がオブジェクトxを記憶

するのに十分となるまで、前記メモリ内の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、

20 オブジェクトxを前記自由空間に記憶するステップとを含む、

請求項1に記載の記憶域空間の節約方法。

【請求項42】もしyが多解像度オブジェクトでなければ、

前記自由空間の量を増加するためにオブジェクトyを前記メモリから除去するステップと、

前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、

30 もし前記自由空間の量がオブジェクトxを記憶するのに十分でなければ、

前記自由空間の量がオブジェクトxを記憶するのに十分となるまで前記メモリ内の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ス

テップを反復するステップと、

オブジェクトxを前記自由空間に記憶するステップとを含む、

請求項41に記載の記憶域空間の節約方法。

【請求項43】確認する前記ステップが前記オブジェクトの順序リストのメモリ・ディレクトリ・エントリを形成するステップを含む、請求項41に記載の記憶域空間の節約方法。

【請求項44】使用近時性最小オブジェクトが多解像度オブジェクトであるかどうかを確認するステップと、

もしオブジェクトyが多解像度オブジェクトであれば、

オブジェクトyの低減解像度版を発生及び記憶するステップと、

自由空間を作成するように、オブジェクトyの現在の高解像度版を前記メモリから除去するステップと、

50 オブジェクトyの優先順位を使用近時性最大オブジェク

トに設定するステップと、
前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、
もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、
オブジェクトxを前記自由空間に記憶するステップとを含む、

請求項 1 に記載の記憶域空間の節約方法。

【請求項 4 5】前記メモリが一時的バッファである、請求項 1 に記載の記憶域空間の節約方法。

【請求項 4 6】前記メモリがキャッシュである、請求項 1 に記載の記憶域空間の節約方法。

【請求項 4 7】コンピュータが多解像度オブジェクトに対する要求を満足させる方法であって、
多解像度オブジェクトを、要求された解像度レベルで渡すべき旨の前記要求を受け取るステップと、
前記コンピュータが、前記要求された解像度レベルの前記オブジェクトを有するかどうかを決定するステップと、

前記コンピュータが、もし前記要求された解像度レベルの前記オブジェクトを有するならば、前記要求されたオブジェクトを渡すステップと、

前記コンピュータが、もし前記要求された解像度よりも高い解像度レベルの前記要求されたオブジェクトを有するならば、

前記要求された解像度レベルの前記オブジェクトを発生するステップと、

前記オブジェクトを渡すステップとを含む、
オブジェクトに対する要求を満足させる方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はデータ処理システム及び、又はコンピュータ・ネットワークにおける記憶メモリに関する。さらに具体的には、本発明はそのようなシステム及び、又はネットワークでデータ・オブジェクトの置換の意思決定を実行する方法及び装置に関する。

【0002】

【従来の技術】メモリ空間の節約及び効率的な使用は記憶ユニットに対して重要である。メモリ空間は一般的なメモリにとって重要な考慮事項であるが、限られた空間の一時的、短期又は長期のメモリ・バッファにとっては最も重要である。例えば、多くのデータ処理システムでは、中央処理ユニットの作業記憶装置と主記憶装置との間には、通常「キャッシュ」と呼ばれる高速メモリ・ユニットが設けられる。このユニットによって、主記憶装置から前もってキャッシュへ転送されたデータのサブセットを比較的高速でアクセスすることができる。通常、主記憶装置はキャッシュよりも低速のメモリであるか、ハード・ディスク記憶装置である。従って、データ処理

システムの動作速度が改善される。同様に、コンピュータ・ネットワークでは、遠隔のネットワーク・コンピュータから前もって転送されたデータへ比較的高速でアクセスするために、キャッシュが使用される。遠隔のネットワーク・コンピュータは一般的に中間プロキシ・コンピュータであるか、元のウェブ・サーバ・コンピュータである。インターネットはそのようなコンピュータ・ネットワークである。

【0003】キャッシュからの情報検索を可能にするため、キャッシュ・エントリーの「ディレクトリ」が維持される。キャッシュに存在する各データ・オブジェクトは、そのアドレスをディレクトリに記憶されている。一度キャッシュがいっぱいになると、1つまたは複数の古いデータ・オブジェクトが除去されるか、本発明の主題のように、1つまたは複数の古いデータ・オブジェクトがサイズを縮減された場合にのみ、新しいデータ・オブジェクトを入れることができる。オブジェクトを置換候補として選択し、キャッシュ内容の変更後にディレクトリを更新するには、ある種のプロシージャが必要である。

【0004】キャッシュ記憶装置を使用し、置換選択及びディレクトリ更新のメカニズムを提供する多数のシステムが当技術分野で知られている。例えば、アール・イー・ランゲ (R. E. Lange) らへの米国特許第4,322,795号は、キャッシュ・ミスのとき主メモリからフェッチされたデータを記憶するキャッシュ位置を選択するために、使用近時性最小 (least recently used すなわち LRU) 方式を使用するキャッシュ・メモリ構成を開示している。

【0005】置換すべきオブジェクトを発見するためには、幾つかの方法が知られている。これは、例えば、1996年11月 (改訂は1997年3月) の IBM リサーチ・レポート (IBM Research Report) 20619 に発表された「ウェブ・オブジェクトのキャッシング方針について」(On Caching Policies for Web Objects) と題するアグガーワル (Aggarwal) らの論文に記載された方法を含む。この論文は参照によりその全体がここに組み込まれる。これらの方法は、使用近時性最小オブジェクトを置換する代わりに、どのデータ・オブジェクトを選択して除去するか の代替方法を提供する。特に、これら方法のいくつかは、置換すべきオブジェクトを選択するときオブジェクト・サイズのようなファクタを考慮する。

【0006】さらに、ジェイ・ロビンソン (J. Robinson) への米国特許第5,043,885号は、特定の方法で維持される参照カウントに基づいて非LRU置換選択を行う方法を開示している。この特許は、参照によりここに組み込まれる。

【0007】これらの方法のいずれも、単一オブジェクトの複数の解像度版について言及しておらず、それを使用してもいない。実際に、置換すべきオブジェクトを発

見するこれらの代替方法は、後に説明するように、本発明の実施例へ直接的な方法で組み込むことができる。

【0008】

【発明が解決しようとする課題】本発明の目的はメモリ空間を節約するための方法、装置、及び製造物に関する。この方法は、多解像度オブジェクトの解像度を動的に低減する。それは、動的に縮減されたオブジェクトを候補置換可能オブジェクトの順序リストへ再統合してそのオブジェクトの複数の解像度レベルを管理するメカニズムを提供する。さらに、1つのコンピュータが他のコンピュータからオブジェクトの所与の解像度レベルを要求することのできるプロシージャが提供される。

【0009】

【課題を解決するための手段】本発明の実施例では、データ・オブジェクトは、通信ネットワークを介して検索されており、通常いくらかの遅延を伴って、プロキシ又はウェブ・ブラウザ又はこれら双方の中間メモリに記憶される。データ・オブジェクトが最初に検索されるとき、前には存在していなかったオブジェクトの低解像度版が計算され、高解像度版と共に保存される。いくつかの場合には、これは非常に小さな追加作業と共に達成される。代替的に、オブジェクトの低解像度版は、前に記憶されていた完全解像度版を削除する前に計算される。この最後の方法は、データ変換が計算的に高価でない場合に特に有用である。典型的には、低解像度版は完全解像度版よりも少ない記憶空間を必要とするであろう。

【0010】この方法を実行する装置では、メモリ・ヒットがデータ・オブジェクトの低解像度版に対して生じたとき、データ・オブジェクトが低解像度であることは、各種の手段によってユーザ又はアプリケーションに知らせることができる。例えば、像が低解像度形式で表現されるウェブ・ブラウズ環境では、特別の囲みボックスのような図形表示が使用される。次に、ユーザは完全解像度版を検索するためにマウス・クリックを実行する。他の環境では、データは低解像度版として明白に示され、同時に縮減の程度が示される。

【0011】さらに、本発明は製造物に記憶することのできる特徴を含む。例えば、製造物はコンピュータへ挿入することのできる1つまたは複数のフロッピー・ディスクの形式を取ることができる。これは、その製造物が使用されるシステムに本発明の属性を与える。

【0012】

【発明の実施の形態】本発明はメモリ空間を節約する方法、装置、製造物に関し、解像度を低減することのできるデータ・オブジェクトに応用することができる。この方法によって、後に説明する実施例で明らかになるように、有効なメモリ・サイズを大きく増加することができる。本発明は一般的に一時的、短期、又は長期の記憶域アプリケーションに応用できる。本発明の1つの例では、データ・オブジェクトがワールド・ワイド・ウェブ

・サーバから通信ネットワークを介して検索され、中間のプロキシ及びクライアントのウェブ・ブラウザにあるキャッシュに記憶されるシナリオを考える。従って、そのようなデータ・オブジェクトをネットワーク・オブジェクト又は単にオブジェクトと呼ぶことにする。

【0013】図1のシステムで、クライアント・コンピュータ100は第2のコンピュータ（プロキシ160）へ接続され、プロキシ160は第3のコンピュータ（ウェブ・サーバ170）へ接続される。典型的なクライアント・サーバ通信では、クライアント・コンピュータ100で実行されているウェブ・ブラウザ・アプリケーションは、ウェブ・サーバ170からプロキシ160を介してネットワーク・オブジェクトを要求する。クライアントの要求に回答して、ウェブ・サーバ170はネットワーク・オブジェクトをプロキシ160へ送り、プロキシ160はそれをクライアント・コンピュータ100へ送る。このシナリオでは、キャッシュはクライアント及びプロキシ、また場合によっては、サーバに存在する。

【0014】サーバ、プロキシ、及びクライアントの各コンピュータのハードウェア及びソフトウェア構成は概して類似している。図2に示されるように、中央処理ユニット（CPU）211及びハード・ディスク記憶装置213のような標準的コンピュータ・システム構成要素のほかに、ネットワーク・オブジェクトの解像度を低減するための特定ハードウェア構成要素212が一般的に含まれている。そのようなハードウェア・ユニットは処理集中性が非常に高い解像度低減プロシージャ（例えば、ビデオ・データのデコード及び解像度低減再エンコード）に使用される。メイン・メモリ216に記憶されるように示されるキャッシュ関連ソフトウェア構成要素は、キャッシュ・ディレクトリ214及びキャッシュ・メモリ217を含み、これらはネットワーク・オブジェクト・キャッシュ215を構成する。キャッシュ・マネージャ220はキャッシュ・ディレクトリ214を更新し、キャッシュ・ディレクトリ内の情報を使用して、ネットワーク・オブジェクトをキャッシュ・メモリ217に記憶し、アクセスし、除去する。これらの構成要素のほかに、さらにソフトウェアのネットワーク・オブジェクト解像度低減プロシージャ218がメモリに記憶されている。データ集中性が高くない解像度低減プロシージャ（例えば、テキスト短縮及び画像縮減の大部分の「タイプ」）は、より典型的にはソフトウェアで実現される。システムの他のソフトウェア構成要素は、外部アプリケーション・プログラム219の集合である。プロキシ・システムでは、これらのプログラムは、ネットワーク・オブジェクトがプロキシを通過するときネットワーク・オブジェクトの解像度を自動的に低減してプロキシとクライアントの間のリンク上で通信帯域幅を節約するようなメカニズムを含んでよい。これは、クライアントがワイヤレス・ネットワークを介して接続される場合で

ある。クライアント・システムでは、アプリケーション・プログラムは典型的にウェブ・ブラウザを含み、ユーザはウェブ・ブラウザを使用してネットワーク・オブジェクトを要求し観察する。

【0015】本発明の目的のためには、「多解像度オブジェクト」の用語は、オブジェクトが、現在表現されている解像度よりも低い解像度（より少ないメモリ空間を使用する）で表現することのできるオブジェクトを意味する。これは、一般的には、オブジェクトが使用されているシステムが低解像度を形成又は獲得する手段を有する条件に依存する。多解像度オブジェクトは多解像度タイプであると呼ばれる。多解像度データ「タイプ」の例、及び解像度低減方法の幾つかの例が説明される。実施例で使用される場合、オブジェクトの低解像度版は、そのオブジェクトの高解像度版よりも少ない記憶域空間を必要とするものと仮定する。

【0016】フォーマットされていないテキスト・オブジェクトは、各種の短縮程度によって縮減される。構造化されたテキスト・オブジェクト（例えば、HTML）は構造的構成要素を除去することによって縮減される。例えば、構造化されたテキスト・オブジェクトの要約構成部分は、そのオブジェクトの低減解像度版として使用される。画像については、より少ない画素、より少ないカラーを有する画像をデコード及び再エンコードする（すなわち、トランスコードする）ことによって、又はより少ないビット／画素を達成する他の方法によって、解像度の低減が達成される。オーディオ・オブジェクトは、低サンプリング・レートへトランスコードするか、サンプルの量子化を粗くすることによって縮減される。ビデオ・オブジェクトは、ビデオ・フレームをドロップするかストリームを完全に再エンコードすることによって縮減される。特定のアプリケーションを最良に満足させるため、前記の任意の組み合わせが使用される。

【0017】図3は、ネットワーク通信スタックの層、及びそれら層とキャッシュ・マネージャの相互作用を示す。ネットワーク・オブジェクト・キャッシュはサーバ、プロキシ、又はクライアントに存在することができる。例示として、図1を参照すれば分かるように、キャッシュがプロキシ160にのみ存在する場合を説明する。クライアント・コンピュータ100上のウェブ・ブラウザが特定のネットワーク・オブジェクトを要求する場合を考える。ネットワーク・オブジェクトはユニバーサル・リソース・ロケータ（Universal Resource Locator、すなわちURL）によって記述され、ブラウザ（又はプロキシ）はURLを使用して要求されたデータを有するウェブ・サーバ170を突き止めることができる。それを行うために、クライアントのウェブ・ブラウザはHTTPを使用して所望のURLを含むHTTP GET要求を出す。HTTPプロトコルは、TCP/IPプロトコルを通信に使用するアプリケーションである。U

RLは、要求されたデータを有するウェブ・サーバ170の宛先アドレスをTCP/IPが獲得することを可能にする十分な情報を含む。これは、TCP/IPプロトコル群のドメイン名システム（DNS）プロトコルを使用して標準方式で達成される。ここでは、中間プロキシが含まれる図1のアーキテクチャを考えているから、最初のステップは、クライアント・コンピュータ100のウェブ・ブラウザ・アプリケーションが前記のHTTP

GET要求をプロキシ160へ渡すことである。プロキシのTCP/IPアドレスは、典型的にユーザによってウェブ・ブラウザへ入れられるパラメータであることに注意されたい。次に、HTTPプロトコルはプロキシへのTCP接続を確立し、要求を送り、それが到着したときプロキシの応答を読み取る。図3を参照すると、HTTPプロトコル層251がTCP/IPプロトコル群252から受け取ったHTTP GET要求をプロキシ・アプリケーション250が受け取ったとき、それは所与のURLをキャッシュ・マネージャ254へ渡す。このURLはネットワーク・オブジェクトを指定しており、キャッシュ・マネージャはそのネットワーク・オブジェクトをキャッシュ・ディレクトリ214の中で探索する。もしキャッシュ・マネージャが要求されたデータを発見すると、キャッシュ・マネージャはそのデータをプロキシ・アプリケーション250へ戻し、プロキシ・アプリケーション250はデータにHTTPヘッダを付けて、それをHTTP応答としてHTTPプロトコル層へ戻し、そのHTTP応答は各種の通信層を介してクライアントのウェブ・ブラウザへ送られる。もしキャッシュ・マネージャが要求されたデータを発見しないと、それはプロキシ・アプリケーション250へ通知し、プロキシ・アプリケーション250はHTTP GET要求をウェブ・サーバ170へ送るが、ソース・アドレスはプロキシ・アドレスへ設定して、応答がプロキシへ戻されるようにする。ウェブ・サーバ170がHTTP GET要求を受け取ったとき、それはHTTP応答をデータ又はエラー・メッセージと共に戻す。ここでは、エラー・メッセージではなく、ネットワーク・オブジェクトが戻されるものと仮定する。プロキシ・アプリケーション250がこのHTTP応答を受け取ったとき、それはキャッシュ・マネージャ254を使用して、それをキャッシュ・メモリ255に記憶するが、さらにプロキシ・アプリケーション250はデータを含むHTTP応答をクライアント・コンピュータ100のウェブ・ブラウザへ送る。

【0018】図4を参照すると、ネットワーク・オブジェクト・キャッシュは、典型的にはネットワーク・オブジェクトのキャッシュ・ディレクトリ・エントリー（CDE）320、330、又は340を有するキャッシュ・ディレクトリ300を使用することによって実現される。CDE 320、330、340は、オブジェクトが

キャッシュ・メモリ310に記憶されている位置を指すポインタを有する。キャッシュ・メモリ310は3つのネットワーク・オブジェクト350、360、370を含むように示される。

【0019】ネットワーク・オブジェクト・キャッシュを管理するためにキャッシュ・マネージャ254によって使用されるCDEの例は、図5に示される。本発明の一部として使用されるCDEの概念は他のアプリケーションでも使用可能であることに注意されたい。この例の目的のためには、ネットワーク・オブジェクト名のフィールド410へ適用されるハッシュ機能を使用することによって、CDEが発見されるものと仮定する。それは、所与のハッシュ値を有するCDE400を指すポインタを提供するか、必要であれば同じハッシュ値を有するCDEの順序リスト（例えば、二重リンク・リスト、すなわちハッシュ・チェーン）をたどる（フィールド425及び440を使用して）。これは、所与のネットワーク・オブジェクト名を有するCDEが発見されるまで継続する。所与の名前を有するオブジェクトのCDEを発見する他の方法が当技術分野で知られている。これらの方法の中には、例えば、すべてのCDEの線形走査、ツリー構造インデックスなどを使用するものがある。これら方法の任意のものを、以下もっと直接的に説明する本発明の実施例で、代替方法として使用することができる。

【0020】図5に戻って、完全なネットワーク・オブジェクト名がフィールド410に記憶され、オブジェクト・タイプ（例えば、HTML、GIF、JPEG、など）がフィールド420に記憶されている。次に、キャッシュ内のオブジェクトがアクセスされた相対的順序が、使用近時性最小（LRU）チェーンを実現する二重リンク・リストを使用して記憶される。これは、使用近時性最大（most recently used、すなわちMRU）オブジェクト及び使用近時性最小オブジェクトのCDEへのポインタ（図示されていない）によって維持される。さらに、二重リンク・リストを実現するために、各CDEはフィールド415及び430を有する。フィールド415は、使用近時性がより大きい次のオブジェクトのCDEを指すポインタを含むか、CDEが使用近時性最大オブジェクトに対するものである場合にはヌル（NULL）を含む。フィールド430は使用近時性がより小さい次のオブジェクトのCDEを指すポインタ（又はヌル）を含む。次に、同じハッシュ値を有するCDEの二重リンク・リストはフィールド440を使用して実現される。フィールド440はハッシュ・チェーン内の前又は次のCDEを指すポインタを含む。もしCDEがハッシュ・チェーン中で最初又は最後のものであれば、各ポインタはヌルを有する。最後に、キャッシュ中のオブジェクトの位置及びサイズがフィールド450に記憶される。

【0021】現在キャッシュ内に存在しない新しいネッ

トワーク・オブジェクトを記憶する空間を作成するために、ネットワーク・オブジェクト・キャッシュ内でオブジェクトを除去する方法は、図6に示される。ステップ510では、オブジェクト'x'がキャッシュに存在せず、'Sx'バイトの記憶域を必要とすることが決定される。ステップ520では、キャッシュ内の現在の自由空間（'S_FREE'で示される）が'Sx'よりも大きいのか、それに等しいかが検査される。もしそうであれば、制御はステップ530へ進み、そこで'x'がキャッシュに記憶される。キャッシュ・メモリの'Sx'バイトが割り振られる。'x'のCDEが作成される（そのCDEをハッシュ・チェーン内でリンクする）。そのCDEは、使用近時性が最大に設定される。'S_FREE'は'S_FREE'-'Sx'へ設定され、'Sx'及び'x'はキャッシュの新しく割り振られたメモリ領域へロードされる。次に、この方法はステップ550で終了する。しかし、ステップ520で、もし存在する自由なキャッシュ・メモリが不十分であると決定されると、制御はステップ540へ進む。このステップで、使用近時性が最小のオブジェクト'y'が発見される。このオブジェクトへ割り振られた'Sy'バイトの空間が解放される。'S_FREE'は'S_FREE'+ 'Sy'へ増加される。'y'はもはやキャッシュに記憶されていないので、'y'のCDEは除去される。次に、制御はステップ520へ戻り、そこで再びxを記憶するための自由なキャッシュ・メモリが十分にあるかどうかを検査される。

【0022】置換すべきオブジェクトを発見する多くの方法が当業者に知られている。非LRU置換選択を実行する前記のような方法の例は、1991年8月27日にジェイ・ロビンソン（J. Robinson）へ与えられた「動的頻度ベース置換及び境界基準を使用するデータ・キャッシュ」（Data Cache Using Dynamic Frequency Based Replacement and Boundary Criteria）と題する米国特許第5,043,885号に説明されている。この特許は、参照してここに組み込まれる。この場合、非LRU選択は特別の方法で維持される参照カウントを使用して行われる。前記の方法のいずれも、単一オブジェクトの複数解像度版に言及しておらず、使用してもいい。さらに、これらの方法は、例えば、1995年にボストンで開催された第四回国際ワールド・ワイド・ウェブ・コンファレンスで発表された、アブラムス（Abrams）らによる論文「キャッシング・プロキシ：制約と可能性」（Caching Proxies: Limitations and Potentials）、1996年のACM SIGCOMM会報（the Proceedings of the ACM SIGCOMM）の293～304頁で発表されたウィリアムス（Williams）らの論文「ワールド・ワイド・ウェブ・ドキュメントのネットワーク・キャッシュにおける除去方針」（Removal Policies in Network Caches for WorldWide Web Documents）、及び1996年11月（改訂は1997年3月）のIBMリサーチ・レポート（IBM Research Report）20619に発表されたアグガーワル（Aggarwal）らの論文「ウ

ウェブ・オブジェクトのキャッシング方針について」(On Caching Policies for Web Objects)に説明される方法を含む。これら論文のすべては、参照してその全体をここに組み込まれる。置換すべきオブジェクトを発見する方法に関しては、その方法はステップ540の代替及び/又は変更として考えられる。使用近時性が最小のオブジェクトを置換する代わりに使用される方法を前に説明した。従って、これらの他の方法では、使用近時性が最小のオブジェクトを置換する代わりに、1つまたは複数の他のオブジェクトが置換可能として発見される。これらの方法の例は、FBR(頻度ベース置換)法、LRU-MIN, SIZE法、及びPSS法を含む。本発明は置換すべきオブジェクトを発見する任意の方法と共に使用される。FBR法を使用するとき、参照カウントを増加することができる。他のオプションは、低減解像度オブジェクトを、例えばLRUチェーンの中間に挿入することを含む。概略的には、本発明の1つの局面は低減解像度オブジェクトの「優先順位」を増加して、そのオブジェクトが即時置換適格オブジェクトにならないようにすることである。ある場合には、オブジェクトのサイズが、置換すべきオブジェクトの選択ファクタとして使用されてよい。実際、置換すべきオブジェクトを発見するこれらの代替方法は、本発明の実施例へ直接的に組み込むことができる。

【0023】本発明の実施例で使用されるCDEの例は、図7に示される。拡張されたCDEは3つの解像度レベルを有する。この設計を異なった数の解像度レベルへ拡張することは簡単であることを認識されたい。この拡張CDEの最初の部分は図5に示されるCDEの例と同じである。図7の拡張CDEは、ネットワーク・オブジェクト名フィールド610、タイプ・フィールド620、使用近時性がより大きい次のオブジェクト及び使用近時性がより小さい次のオブジェクトのCDEを指すポインタ630、及びハッシュ・チェーンにおける前及び次のCDEを指すポインタ640を含む。3つのフィールド650、660、670はキャッシュ内のオブジェクトの完全解像度版及び低減解像度版の位置及びサイズを与える。これらフィールドの任意のものがヌルであってよく、そのヌルは、オブジェクトの版が所与の解像度で現在キャッシュに存在しないことを示す。しかし、すべてのフィールドが同時にヌルであることはできない。同時のヌルは、オブジェクトの版がキャッシュに存在しないことを示す。そのような場合、CDEはキャッシュ・ディレクトリから除去される。詳細に説明すれば、フィールド650はオブジェクトの完全解像度版の位置及びサイズを与える。完全解像度版は、ここでは解像度「レベル0」と呼ばれる。フィールド660は、オブジェクトのタイプに関連づけられた解像度低減プロシーダを1回適用することによって得られた、オブジェクトの低減解像度版の位置及びサイズを与える。この低減解

像度は、ここでは解像度「レベル1」と呼ばれる。フィールド670は、オブジェクトの更に低減された解像度版の位置及びサイズを与える。その低減解像度は、ここでは解像度「レベル2」と呼ばれる。「レベル2」は、オブジェクトの完全解像度版に対して、オブジェクトのタイプに関連づけられた解像度低減プロシーダを2回適用することによって得られる。又は、オブジェクトの解像度「レベル1」版に対して1回適用することによっても同様に得ることができる。

10 【0024】本発明の実施例は、ここでは2つの異なったケースに分けて説明される。ケース1は、一時にオブジェクトの1つだけの版(すなわち、解像度レベル)がキャッシュに存在する場合である。ケース2は、オブジェクトの複数の解像度レベルがキャッシュにあらかじめ存在することができる場合である。ケース1は、今日のインターネットで通常使用される一般的なウェブ・キャッシュの典型である。ケース2の状態は、特殊タイプの中間プロキシ(図1の160を参照)が使用される場合である。トランスコーディング・プロキシと呼ばれるこの特殊プロキシは、サーバから送られたネットワーク・オブジェクトを傍受し、オブジェクトの低解像度版を形成する。次に、この低解像度版は高解像度版と共にプロキシ・キャッシュに記憶される。ケース2では、クライアント・リンクへのプロキシ上の通信帯域幅を節約するために、オブジェクトの高解像度版ではなく、低解像度版がクライアントへ送られる。プロキシ内に記憶されたオブジェクトの高解像度版は、それが必要であることをクライアント(図1の100を参照)が決定すれば、要求に応じて利用可能である。

30 【0025】図8は、ケース1の場合の、本発明の実施例を示す。それは、現在キャッシュに存在しない新しいネットワーク・オブジェクトを記憶する空間を作成するために、ネットワーク・オブジェクト・キャッシュ内のオブジェクトを管理し、追加し、除去し、又は縮減する方法を示す。680、681、682、及び683はキャッシュ・ディレクトリ・エントリ(CDE)であり、それらは図7に示されるようなネットワーク・オブジェクトを指すポインタを有する。これらのCDEは、使用近時性が最大のオブジェクトから始まって使用近時性が最小のオブジェクトで終わるリンク・リストに記憶される。ここでは、このリンク・リストをLRUチェーンと呼ぶことにする。図8に示される状況では、新しいオブジェクトを記憶するために追加の空間がキャッシュ内に必要であると決定された場合、使用近時性が最小のオブジェクトが検査される。この使用近時性最小オブジェクトはCDE683によって指し示される。もしこのLRUオブジェクトが多解像度タイプでなければ、それはキャッシュ内に追加の空間を作るために除去される。もしオブジェクト684が多解像度タイプであれば、低解像度版685が発生され、この低解像度版を指し示す

CDEが前記LRUチェーンの始めに挿入される。オブジェクトの高解像度版はキャッシュ内に追加の空間を作るために廃棄される。このようにして、ケース1の場合、ネットワーク・オブジェクトは、新しいオブジェクトのためにキャッシュ内で空間をクリアするため、動的にサイズを縮減されることができる。

【0026】現在キャッシュに存在しない新しいネットワーク・オブジェクトを記憶する空間を作成するためにネットワーク・オブジェクト・キャッシュ内でオブジェクトを除去又は縮減する方法は、図9に示される。図9は、前述したケース2の、本発明の実施例を示す。690、691、692、及び693は図8に示されるようなLRUチェーンを形成するキャッシュ・ディレクトリ・エントリーである。

【0027】図9は、新しいオブジェクトを記憶するため、追加の空間がキャッシュ内で必要であると決定される場合のステップを示す。CDE 693によって指し示される使用近時性が最小のオブジェクトが検査される。もしこのLRUオブジェクトが多解像度タイプでなければ、それは追加の空間をキャッシュ内に作るために除去される。もしオブジェクト694が多解像度タイプであれば、ケース2の場合には複数の解像度版が存在してもよいので、現存する最高解像度版オブジェクト694が廃棄され、そのネットワーク・オブジェクトのCDEが前述したLRUチェーンの始めに挿入される。このようにして、ケース2では、新しいオブジェクトのためにキャッシュ内の空間をクリアするため、ネットワーク・オブジェクトの複数の解像度版の中で最大の版が廃棄される。

【0028】前述したように、高レベルで、ケース1及びケース2に対応する本発明の実施例の方法が、2つの詳細なフローチャートで図10及び図11に示される。本発明の実施例で、現在キャッシュに存在しない新しいネットワーク・オブジェクトを記憶する空間を作るためにネットワーク・オブジェクト・キャッシュ内のオブジェクトを除去及び縮減する方法は、図10に示される。

【0029】ケース1の場合に、本発明の実施例でネットワーク・オブジェクト・キャッシュ内のオブジェクトを除去又は縮減する方法は、図10に示される。ステップ706、711、716、及び721は、前に説明した図6の方法に対応する。ステップ706では、オブジェクト'x'がキャッシュに存在しないこと、及び'Sx'バイトの記憶域を必要とすることが決定される。ステップ711では、'S_FREE'で表されたキャッシュ内の現在の自由空間が'Sx'よりも大きいか、それに等しいかが検査される。もしそれが大きければ、制御はステップ716へ進む。ステップ716で、'x'はキャッシュに記憶され、'Sx'バイトがキャッシュ・メモリで割り振られる。'x'のCDEが作成されて、そのCDEがハッシュ・チェーンにリンクされる。そのCDEは使用近時性が

最大であると設定され、'S_FREE'が'S_FREE'-'Sx'へ設定される。オブジェクト'x'はキャッシュの新しく割り振られたメモリ領域にロードされる。次に、この方法はステップ721で終了する。しかし、ステップ711で、もし存在する自由なキャッシュ・メモリが不十分であると決定されると、制御はステップ731へ進む。ステップ731で、前述したように、使用近時性が最小のオブジェクト'y'が発見されるか、他の置換候補オブジェクトが発見される。この時点で、前の方法とは異なり、ステップ736で、次のようにして'y'が多解像度オブジェクトであるかどうか検査される。システムは、解像度低減プロシージャが利用できるネットワーク・オブジェクト「タイプ」のリストを維持している。図2に示されるように、それらの解像度低減プロシージャはハードウェア又はソフトウェアで実現することに注意されたい。もし'y'の「タイプ」がこのリストにあれば、ステップ736のテストが進行し、'y'が多解像度オブジェクトであることが示される。そうでなければ、'y'はステップ741に従って処理される。'Sy'バイトから構成される'y'へ割り振られた空間は解放され、'S_FREE'は'S_FREE'+'Sy'へ増加される。'y'はもはやキャッシュに記憶されていないので、'y'のCDEは除去される。次に、制御はステップ711へ戻り、そこで'x'を記憶するための自由なキャッシュ・メモリが十分に存在するかどうか検査される。

【0030】しかし、ステップ736で、もし'y'が多解像度オブジェクトであることが発見されると、制御はステップ751へ進む。ステップ751で、オブジェクト'y'の解像度「レベル0」（完全解像度）版がキャッシュに存在するかどうか検査される。もしそれが存在すれば、解像度「レベル0」位置フィールドはヌルではなく、制御はステップ755へ進む。ステップ755で、'y'の「タイプ」に関連づけられた解像度低減プロシージャが'y'の解像度「レベル0」版へ適用され、'y'の解像度「レベル1」版が発生される。'y'_RES_0の空間がキャッシュ内で解放され、解像度「レベル0」位置がヌルに設定される。解像度レベル版'y'_RES_1がキャッシュに記憶され、CDEの解像度1位置フィールドが、その位置を指し示すために設定される。解像度「レベル0」版は除去されたので、'S_FREE'が最初に'S_FREE'+'Sy'_SIZE_0へ増加される。解像度「レベル1」版が記憶されたので、'S_FREE'は'S_FREE'-'Sy'_SIZE_1へ減少される。'y'のCDEは使用近時性が最大に設定され、制御はステップ711へ戻る。

【0031】ステップ751で、もし解像度「レベル0」版が発見されないと（すなわち、位置フィールドはヌルであった）、制御はステップ761へ進む。ステップ761で、'y'の解像度「レベル1」版がキャッシュに存在するかどうか検査される。もし存在すれば、制御はステップ765へ進む。ステップ765で、'y'の

「タイプ」に関連づけられた解像度低減プロシージャが'y'の解像度「レベル 1」版へ適用され、サイズ'Sy'_SIZE_2を有する'y'の解像度「レベル 2」版が発生される。'y'_RES_1の空間がキャッシュ内で解放され、解像度「レベル 1」位置がヌルに設定される。解像度「レベル 2」版の'y'_RES_2がキャッシュに記憶され、CDEの解像度 2 位置フィールドは、その位置を指し示すように設定される。最初に、'S_FREE'が'S_FREE'+'Sy'_SIZE_1へ増加され、解像度「レベル 1」版が除去される。次に、解像度「レベル 2」版が記憶されたので、'S_FREE'が'S_FREE'-'Sy'_SIZE_2へ減少される。'y'のCDEは使用近時性が最大へ設定され、制御はステップ 7 1 1 へ戻る。

【0032】ステップ 7 6 1 へ戻って、もし解像度「レベル 1」版がキャッシュ内で発見されないと、それは解像度「レベル 2」版が存在しなければならないことを示す。それは、この制御通路に沿って、解像度「レベル 0」版が存在しないことが前に決定されたからである。前述したように、CDEが存在するためには、オブジェクトは少なくとも 1 つの解像度レベルでキャッシュに存在しなければならない。その場合、制御はステップ 7 7 1 へ進む。例示目的のために使用される実施例では、2 よりも高い解像度レベルは使用されないものと仮定する。従って、ステップ 7 7 1 で、キャッシュ内のオブジェクトの唯一の版は解像度「レベル 2」で存在するから、解像度の更なる低減は許されず、オブジェクトはキャッシュから除去される。'Sy'_SIZE_2のサイズを有する'y'の解像度「レベル 2」版の空間は解放され、'S_FREE'は'S_FREE'+'Sy'_SIZE_2へ増加される。'y'のCDEはキャッシュ・ディレクトリから除去され、制御はステップ 7 1 1 へ戻される。この設計を拡張して、最大解像度レベルがすべてのオブジェクトについて固定されないようにすることは簡単であることに注意されたい。例えば、最大解像度レベルはオブジェクトの「タイプ」、その「タイプ」のオブジェクトのサイズに依存することができ、解像度のレベルが増加すると、オブジェクトのサイズを増大させることができる。

【0033】前記の方法では、ステップ 7 5 5 及び 7 6 5 で説明したように、低減解像度オブジェクトのCDEは使用近時性が最大に設定された。低減解像度オブジェクトをCDEのリンク・リストへ再統合するために、代替の方法を使用してもよい。例えば、1 つのそのような代替方法では、低減解像度オブジェクトはCDEのリンク・リストの中間に再挿入される。他の代替方法は、ジェイ・ロビンソン (J. Robinson) らへの米国特許第 5,043,885 号で説明されている参照カウント・ベース法を使用することである。それは、例えば、リンク・リスト位置を維持するが、参照カウントを増加することによって行われる。これら代替方法のすべては、低減解像度オブジェクトの優先順位を増大させる効果を有し、そのオブ

ジェクトが置換又は更なる縮減のために即時に選択されないようにする。

【0034】ケース 2 の場合に、本発明の実施例でネットワーク・オブジェクト・キャッシュ内のオブジェクトを除去又は縮減する方法は、図 1 1 に示される。ステップ 7 0 5、7 1 0、7 1 5、及び 7 2 0 は前に図 6 で示した方法に対応する。ステップ 7 0 5 で、オブジェクト'x'がキャッシュに存在しないこと、及び'Sx'バイトの記憶域を必要とすることが決定される。ステップ 7 1 0 で、'S_FREE'で示されるキャッシュ内の現在の自由空間が'Sx'よりも大きいのか、それに等しいかが検査される。もしそれが大きければ、制御はステップ 7 1 5 へ進む。ステップ 7 1 5 で、'x'がキャッシュに記憶され、'Sx'バイトがキャッシュ・メモリで割り振られる。'x'のCDEが作られ、そのCDEはハッシュ・チェーン内でリンクされる。そのCDEは使用近時性が最大に設定され、'S_FREE'は'S_FREE'-'Sx'へ設定される。オブジェクト'x'はキャッシュの新しく割り振られたメモリ領域へロードされる。次に、この方法はステップ 7 2 0 で終了する。しかし、ステップ 7 1 0 で、もし存在する自由なキャッシュ・メモリが不十分であることが決定されると、制御はステップ 7 3 0 へ進む。前述したように、ステップ 7 3 0 で、使用近時性が最小のオブジェクト'y'が発見されるか、他の置換候補オブジェクトが発見される。この時点で、前の方法と異なり、ステップ 7 3 5 で、次のようにして、'y'が多解像度オブジェクトであるかどうかを検査される。システムは、解像度低減プロシージャを利用できるネットワーク・オブジェクト「タイプ」のリストを維持している。図 2 に示されるように、これらの解像度低減プロシージャはハードウェア又はソフトウェアで実現することができる。もし'y'の「タイプ」がこのリストにあれば、ステップ 7 3 5 のテストが進行し、'y'が多解像度オブジェクトであることが示される。そうでなければ、'y'はステップ 7 4 0 に従って処理される。'Sy'バイトから構成される'y'へ割り振られた空間が解放され、'S_FREE'は'S_FREE'+'Sy'へ増加される。'y'はもはやキャッシュに記憶されていないので、'y'のCDEは除去される。次に、制御はステップ 7 1 0 へ戻り、そこで'x'を記憶するための自由なキャッシュ・メモリが十分に存在するかどうかを検査される。

【0035】しかし、ステップ 7 3 5 で、もし'y'が多解像度オブジェクトであることが発見されると、制御はステップ 7 5 0 へ進む。ステップ 7 5 0 では、オブジェクトの解像度「レベル 0」（完全解像度）版がキャッシュに存在するかどうかを決定するため、オブジェクト'y'が検査される。もしそれが存在すれば、解像度「レベル 0」位置フィールドはヌルではなく、制御はステップ 7 5 2 へ進む。ステップ 7 5 2 では、'y'の解像度「レベル 1」版がキャッシュに存在するかどうかを検査され

る。もしそれが存在すれば、制御はステップ 7 5 6 へ進む。ステップ 7 5 6 では、解像度「レベル 0」版の空間が解放され、解像度「レベル 0」版の'y'の C D E 内の位置がヌルに設定される。'S_FREE'は'S_FREE'+'Sy'_SIZE_0へ増加され、'y'の C D E は使用近時性が最大に設定され、制御はステップ 7 1 0 に戻る。

【0036】ステップ 7 5 2 で、もし解像度 1 版が発見されないと、制御はステップ 7 5 4 へ進む。ステップ 7 5 4 では、'y'の「タイプ」に関連づけられた解像度低減プロシージャが、'y'の解像度「レベル 0」版へ適用され、'y'の解像度「レベル 1」版が発生される。'y'_RES_0の空間がキャッシュ内で解放され、解像度「レベル 0」位置がヌルに設定される。解像度レベル版'y'_RES_1がキャッシュに記憶され、C D E の解像度 1 位置フィールドがその位置を指し示すように設定される。解像度「レベル 0」版は除去されたので、'S_FREE'が最初に'S_FREE'+'Sy'_SIZE_0へ増加される。解像度「レベル 1」版が記憶されたので、'S_FREE'は'S_FREE'-'Sy'_SIZE_1へ減少される。'y'の C D E は使用近時性が最大に設定され、制御はステップ 7 1 0 へ戻る。

【0037】ステップ 7 5 0 で、もし解像度「レベル 0」版が発見されないと（すなわち、位置フィールドはヌル）、制御はステップ 7 6 0 へ進む。ステップ 7 6 0 で、'y'の解像度「レベル 1」版がキャッシュに存在するかどうかを検査される。もし存在すれば、制御はステップ 7 6 2 へ進む。そこでは解像度「レベル 2」版がキャッシュに存在するかどうかを検査される。もし解像度「レベル 2」版が存在すれば、制御はステップ 7 6 6 へ進む。ステップ 7 6 6 では、解像度「レベル 1」版の空間が解放され、解像度「レベル 1」版の'y'の C D E 内の位置がヌルへ設定される。'S_FREE'は'S_FREE'+'Sy'_SIZE_1へ増加され、'y'の C D E は使用近時性が最大に設定され、制御はステップ 7 1 0 へ戻る。

【0038】ステップ 7 6 2 で、もし解像度 2 版が発見されないと、制御はステップ 7 6 4 へ進む。ステップ 7 6 4 で、'y'の「タイプ」に関連づけられた解像度低減プロシージャが、'y'の解像度「レベル 1」版へ適用され、サイズ'Sy'_SIZE_2を有する'y'の解像度「レベル 2」版が発生される。'y'_RES_1の空間がキャッシュ内で解放され、解像度「レベル 1」位置がヌルへ設定される。解像度「レベル 2」版の'y'_RES_2がキャッシュに記憶され、C D E 内の解像度 2 位置フィールドはその位置を指し示すように設定される。最初に、'S_FREE'が'S_FREE'+'Sy'_SIZE_1へ増加され、解像度「レベル 1」版が除去される。次に、解像度「レベル 2」版が記憶されたので、'S_FREE'が'S_FREE'-'Sy'_SIZE_2へ減少される。'y'の C D E は使用近時性が最大に設定され、制御はステップ 7 1 0 へ戻る。

【0039】ステップ 7 6 0 へ戻って、もし解像度「レベル 1」版がキャッシュ内で発見されないと、それは解

像度「レベル 2」版が存在しなければならないことを示す。それは、この制御通路に沿って、解像度「レベル 0」版が存在しないことが前に決定されたからである。前述したように、C D E が存在するためには、オブジェクトは少なくとも 1 つの解像度レベルでキャッシュに存在しなければならない。この場合、制御はステップ 7 7 0 へ進む。例示目的のために使用される実施例では、2 よりも大きい解像度レベルは使用されないものと仮定する。従って、ステップ 7 7 0 では、キャッシュ内のオブジェクトの唯一の版は解像度「レベル 2」であるから、解像度の更なる低減は許されず、オブジェクトはキャッシュから除去される。'Sy'_SIZE_2のサイズを有する'y'の解像度「レベル 2」版の空間は解放され、'S_FREE'は'S_FREE'+'Sy'_SIZE_2へ増加される。'y'の C D E はキャッシュ・ディレクトリから除去され、制御はステップ 7 1 0 へ戻る。この設計を拡張して、最大解像度レベルがすべてのオブジェクトに固定されないようにすることは簡単であることに注意されたい。例えば、最大解像度レベルはオブジェクトの「タイプ」、その「タイプ」のオブジェクトのサイズに依存することができ、オブジェクト・サイズを増大させるために、解像度のレベルを増加することが許される。

【0040】前記の方法で、ステップ 7 5 4、7 6 4、7 5 6、及び 7 6 6 で説明したように、低減解像度オブジェクトの C D E は使用近時性が最大のオブジェクトとして設定された。図 1 0 で前に説明したように、低減解像度オブジェクトを C D E のリンク・リストへ再統合するために、代替の方法を使用してよい。

【0041】オブジェクトが多解像度「タイプ」であるかどうかを決定するとき、図 1 2 に示されるようなテーブルが使用されてよい。このテーブル・ヘッダ 8 1 0 には、多解像度「タイプ」の現在の数（この例では 3）が、実現形態の詳細度に依存して他の情報と共に記憶されてよい。次に、多解像度「タイプ」の現在の集合の各メンバーについて、1 つのエントリーがテーブルに存在する。そのエントリーは、「タイプ」名、及びその「タイプ」に関連づけられた解像度低減プロシージャを指すポインタを与える。それは図 1 2 のエントリー 8 2 0、8 3 0、及び 8 4 0 によって示される。図 2 に示されたように、これらの解像度低減プロシージャは、新しい解像度低減プロシージャが、現在テーブルに存在していない「タイプ」のために利用可能となったとき、動的に拡張することができる。例えば、オブジェクトの 1 つの「タイプ」が「フォーマットされていないテキスト」であり、この「タイプ」は現在多解像度タイプではないと想定する。ある場合には、フォーマットされていないテキスト・ファイルの一種の解像度低減プロシージャとして、例えばテキスト・ファイルの最後の 9 0 % を消去し最初の 1 0 % を残すテキスト短縮プロシージャを考慮することが許されよう。ファイルが短縮されていることを

示すために、縮減されたファイルの終わりヘストリング MORE を付加することができる。そのような場合、図 1 2 の多解像度オブジェクト・テーブルは、次のようにしてこの「タイプ」を含むように拡張することができる。すなわち、エントリー「フォーマットされていないテキスト、短縮テキスト」がテーブルへ付加され、次にテーブル・ヘッダ 8 1 0 内の「タイプ」の数が 4 へ増加される。

【0042】本発明の 1 つの実施例の最後の構成要素メカニズムは、プロキシ・コンピュータのためのクライアントが、他のコンピュータ（すなわち、プロキシ又はサーバ）へ、所与の名前を有する多解像度オブジェクトを所与の解像度レベルで送る要求を指定する方法である。クライアントのブラウザは、オブジェクト名（例えば、Y）及び所望のオブジェクト解像度レベル（例えば、SIZE₁）が要求されるメッセージをソース・コンピュータ（すなわち、プロキシ又はサーバ）へ送る（例えば、HTTP GET 要求）。ソース・コンピュータでは、もし要求された解像度レベルを有するオブジェクトが利用可能であれば、それがクライアントへ送られる。もし要求された解像度を有するオブジェクトが利用可能でなければ、ソース・コンピュータは、要求された解像度レベルを発生するか、他のソースへ要求を行うか、要求されたオブジェクトは要求された解像度レベルでは利用できないことを示すメッセージを戻す。

【0043】さらに、本発明の実施例のすべては、製造物に組み込むのに適している。製造物は本発明を特定のシステムで実現するために使用される。例えば、製造物は、システムに本発明の実施例を実行させるためにロードされるフロッピー・ディスク又は CDROM であってよい。

【0044】他の実施例では、オブジェクト・キャッシュ内のオブジェクトを除去又は縮減する方法は、現在キャッシュに存在しない新しいオブジェクトを記憶する空間の量を作成するステップと、多解像度タイプ・オブジェクトの位置を突き止めることによってキャッシュ内の特定のオブジェクトを置換のために選択するステップと、多解像度タイプのオブジェクトの最高解像度版を廃棄するステップと、オブジェクトの次の最低解像度版が存在するときそれを保持するステップと、多解像度オブジェクトの次の最低解像度版の優先順位を増加するステップと、多解像度オブジェクトが次の置換の即時適格となるのを防止するステップとを含む。オブジェクトの優先順位を増加する多くの方法が、当業者に知られている。それらの例には、オブジェクトの優先順位を増加する基準として使用頻度を使用する FRB（頻度ベース置換）法がある。さらに、優先順位を増加する方法は、LRU-MIN、SIZE 法及び PSS 法で使われる基準に基づかせることができる。

【0045】本発明の説明は特定のステップ構成として

なされたが、本発明の意図及び概念は他の構成にも適しており適用可能である。従って、例えば、製造物は第 1 解像度のデータを有する特定のオブジェクト群を含むことができ、オブジェクトの 1 つの第 1 解像度のデータが、第 1 解像度削除方針に従って、いつ削除されるかをコンピュータに決定させたり、またオブジェクト群の各々が低解像度のデータを有して、前記オブジェクトの 1 つの低解像度データが、低解像度削除方針に従って、いつ削除されるかを決定するアクションをコンピュータに実行させることができる。説明された CDE 及びリンク・リストの使用は、当業者に知られたリストを使用する 1 つの方法にすぎない。従って、任意の順序リストを使用することができる。特に、キャッシュの使用は、本発明の概念を実現する便宜的な例にすぎない。本発明は、記憶ユニットのタイプがどのようなものであれ、空間節約及び／又はメモリ節約の方法、装置、又は製造物を提供することを意図している。開示された実施例に対しては、本発明の趣旨及び範囲を逸脱することなく、他の変更を加えることができることは、当業者に明らかであろう。

【0046】まとめとして、本発明の構成に関して以下の事項を開示する。

（1）複数のオブジェクトをメモリに記憶するユニットで記憶域空間の量を節約する方法であって、オブジェクトの 1 つの第 1 の解像度版をメモリに有するステップと、前記オブジェクトの 1 つの少なくとも 1 つの低解像度版であって、前記第 1 の解像度版よりも少ないメモリ空間を使用する解像度版を発生するステップと、前記低解像度版をメモリに保存するステップと、前記第 1 の解像度版をメモリから廃棄するステップとを含む、記憶域空間の節約方法。

（2）複数のオブジェクトをキャッシュ内で管理する方法であって、存在する非多解像度オブジェクト及び少なくとも 1 つの多解像度オブジェクトのためにキャッシュ・ディレクトリ・エントリーを有するキャッシュ・ディレクトリを作成し、前記多解像度オブジェクトの各キャッシュ・ディレクトリ・エントリーがそのオブジェクトのすべての存在する解像度版を指すポインタを有するようにするステップと、前記エントリーの順序リストを形成するステップとを含む、オブジェクトの管理方法。

（3）前記形成するステップが、特定のオブジェクトが使用された最後の使用時点に基づく、上記（2）に記載のオブジェクトの管理方法。

（4）前記順序リストがオブジェクト使用基準の頻度に従って形成される、上記（2）に記載のオブジェクトの管理方法。

（5）各オブジェクトを現在の解像度版で記憶するステップと、前記順序リスト内の最低優先順位オブジェクト y に対応するキャッシュ・ディレクトリ・エントリーが、多解像度オブジェクトであるかどうかを確認するス

テップと、オブジェクトyが多解像度オブジェクトであれば、オブジェクトyの低減解像度版を発生及び記憶するステップと、前記キャッシュ中に自由空間を作成するためにオブジェクトyの前記現在の解像度版を除去するステップと、オブジェクトyの前記キャッシュ・ディレクトリ・エントリを最高優先順位オブジェクトに設定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リスト中の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、前記自由空間にオブジェクトxを記憶するステップとを含む、上記(2)に記載のオブジェクトの管理方法。

(6) yが多解像度オブジェクトでなければ、前記自由空間の量を増加するためにオブジェクトyを前記キャッシュから除去するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リスト内の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクトxを前記自由空間に記憶するステップとを含む、上記(5)に記載のオブジェクトの管理方法。

(7) 各オブジェクトを現在の解像度版で記憶するステップと、前記順序リストの最後に置かれた使用近時性最小オブジェクトに対応する第1のキャッシュ・ディレクトリ・エントリが多解像度オブジェクトであるかどうかを確認するステップと、オブジェクトyが多解像度オブジェクトであれば、オブジェクトyの低減解像度版を発生及び記憶するステップと、自由空間を作成するように、オブジェクトyの現在の高解像度版を除去するステップと、オブジェクトyの前記第1のキャッシュ・ディレクトリ・エントリを前記順序リストの始めに配置することによって前記第1のキャッシュ・ディレクトリ・エントリを使用近時性最大オブジェクトに設定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリの前記順序リストにおける次の使用近時性最小オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクトxを前記自由空間に記憶するステップとを含む、上記(2)に記載のオブジェクトの管理方法。

(8) yが多解像度オブジェクトでなければ、前記自由

空間の量を増加するためオブジェクトyを前記キャッシュから除去するステップと、前記自由空間がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、前記自由空間の量が十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまでキャッシュ・ディレクトリ・エントリの前記順序リストにおける次の使用近時性最小オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクトxを前記自由空間に記憶するステップとを含む、上記(7)に記載のオブジェクトの管理方法。

(9) 前記管理するステップが、前記多解像度オブジェクトの少なくとも1つのオブジェクト解像度を低減するステップを含む、上記(2)に記載のオブジェクトの管理方法。

(10) オブジェクト・キャッシュ内のオブジェクトを管理する方法であって、現在前記キャッシュ中に存在しない新しいオブジェクトを記憶するために空間を作成するステップと、多解像度オブジェクトの位置を決定することによって前記キャッシュ中の特定のオブジェクトを置換のために選択するステップと、前記多解像度オブジェクトの最高解像度版を廃棄するステップと、前記オブジェクトの次の最低解像度版があれば、それを保持するステップと、前記多解像度オブジェクトの次の最低解像度版の優先順位を増加するステップと、前記多解像度オブジェクトが後の置換に対して即時に適格となることを防止するステップとを含む、オブジェクトの管理方法。

(11) 増加する前記ステップが、特定のオブジェクトが使用された最後の使用時点に基づいて順序リストを形成するステップを含む、上記(10)に記載のオブジェクトの管理方法。

(12) 複数のオブジェクトをキャッシュに記憶するシステムでオブジェクトが多解像度タイプであるかどうかを決定する方法であって、前記システム中で前記現在理解された多解像度タイプ及び各オブジェクト・タイプのために低減解像度版を発生する前記プロシージャをリストするテーブルを作成するステップと、オブジェクトの前記タイプと前記テーブル中にリストされた複数のタイプの1つとの間で一致を探索するステップと、一致が発見されなければ、前記オブジェクトが多解像度オブジェクトではないと表示するステップと、一致が発見されれば、前記オブジェクトが多解像度オブジェクトであること、及び前記テーブル中にリストされた解像度低減プロシージャが前記オブジェクトの低解像度版を発生するために使用できることを表示するステップとを含む、多解像度オブジェクトの決定方法。

(13) 第1のコンピュータ・システムが、所与の名前を有する多解像度オブジェクトを所与の解像度レベルで渡すように第2のコンピュータ・システムへ要求する方法であって、前記第1のコンピュータ・システムから前

10

20

30

40

50

記第 2 のコンピュータ・システムへ、前記多解像度オブジェクトを要求するメッセージを送るステップと、前記第 2 のコンピュータ・システムが、前記要求されたオブジェクトを、要求された解像度レベルで有するとき、前記要求されたオブジェクトを前記第 1 のコンピュータ・システムへ渡すステップと、前記第 2 のコンピュータ・システムが、前記要求されたオブジェクトを、高解像度レベルで有し、また前記要求された解像度レベルを発生するのに必要な前記プロシージャを有するとき、前記要求されたオブジェクトを前記第 1 のコンピュータ・システムへ渡すステップと、前記第 2 のコンピュータ・システムが、前記要求されたオブジェクトを、前記要求された解像度レベルで有しないで、前記オブジェクトが存在するとき、前記要求された解像度を発生する方法を有しないとき、前記要求されたオブジェクトが利用できないことを示すメッセージを前記第 1 のコンピュータ・システムへ送るステップとを含む、オブジェクトの要求方法。

(14) 現在前記キャッシュ中に存在しない新しいオブジェクトxを記憶する自由空間を作成するために、オブジェクト・キャッシュ内のオブジェクトを管理する方法であって、各多解像度オブジェクトを、低解像度版によって占められるキャッシュ空間の量が高解像度版によって占められる前記空間の量よりも小さくなるような複数の解像度版で記憶するステップと、存在する非多解像度オブジェクト及び前記少なくとも 1 つの多解像度オブジェクトのためにエントリーを有するキャッシュ・ディレクトリを作成し、各キャッシュ・ディレクトリ・エントリーが前記少なくとも 1 つの多解像度オブジェクトのすべての存在する解像度版を指すポインタを有するようにするステップと、特定のオブジェクトが使用された最後の使用時点に基づく前記順序基準に従って前記エントリーの順序リストを形成ステップと、前記順序リストの前記終わりにある使用近時性最小オブジェクトyに対応する前記キャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、yが多解像度オブジェクトでないとき、オブジェクトyを除去して、新しく増加した自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを確認するステップと、新しく増加した前記自由空間の量がオブジェクトxを記憶するのに十分でないとき、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リスト内の次の使用近時性最小オブジェクトのために決定及び確認の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、yが多解像度オブジェクトであるとき、前記オブジェクトの複数の解像度版が現在記憶されているかどうかを決定するステップと、1 つの解像度版のみが現在記憶されているとき、オブジェクトyの低減解像度版を発生及び記憶するステップと、オブジ

ェクトyの前記現在の高解像度版を除去するステップと、前記オブジェクトyのキャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリーを使用近時性最大オブジェクトに設定し、前記新しく増加した自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを確認するステップと、新しく増加した前記自由空間の量がオブジェクトxを記憶するのに十分でないとき、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リストにある前記次の使用近時性最小オブジェクトのために決定及び確認の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、複数の解像度版が現在記憶されていれば、オブジェクトyの前記現在の高解像度版を除去するステップと、オブジェクトyの前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリーを前記使用近時性最大オブジェクトに設定し、前記新しく増加した自由空間がオブジェクトxを記憶するのに十分であるかどうかを確認するステップと、十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リストにあるオブジェクトyの前記次の最高解像度版のために除去、設定、及び確認の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、オブジェクトの管理方法。

(15) 複数のオブジェクトをキャッシュに記憶するシステムでオブジェクトが多解像度タイプであるかどうかを決定する方法であって、前記現在理解されている多解像度タイプをリストするテーブルを前記システム中に作成するステップと、オブジェクトの前記タイプと前記テーブルにリストされた前記タイプとの間の一致を探索するステップと、一致が発見されないとき、前記オブジェクトが多解像度オブジェクトではないことを表示するステップと、一致が発見されたとき、前記オブジェクトが多解像度オブジェクトであることを表示するステップとを含む、オブジェクトの決定方法。

(16) 前記テーブルが、各オブジェクト・タイプのために低減解像度版を発生するプロシージャをリストしているとき、前記一致が発見されると、前記テーブル内の情報から、前記オブジェクトの低解像度版を発生するために特定の解像度低減プロシージャが使用されるべきであることを表示するステップを含む、上記(15)に記載のオブジェクトの決定方法。

(17) ネットワーク・オブジェクト・キャッシュ内のオブジェクトを変更する方法であって、現在前記キャッシュ中に存在しないオブジェクト'x'によって必要とされる記憶域の前記キャッシュ・サイズ'Sx'バイトを決定するステップと、前記キャッシュ内の現在の自由空間の

量 'S_FREE' が 'Sx' よりも大きいのか、それに等しいかどうかを決定するステップと、'S_FREE' が 'Sx' よりも大きい場合、さらにキャッシュ・メモリの 'Sx' バイトを 'x' へ割り振り、'x' を前記割り振られたキャッシュに記憶するステップと、'x' のためにキャッシュ・ディレクトリ・エントリ（CDE）を作成し、前記CDEをハッシュ・チェーンに配列するステップと、前記CDEを現在の使用近時性最大オブジェクトに設定するステップと、'S_FREE' を 'S_FREE' - 'Sx' に設定するステップとを含む、オブジェクトの変更方法。

（18）'Sx' が 'S_FREE' よりも大きいとき、前の使用近時性最小オブジェクト 'y' を発見するステップと、オブジェクト 'y' が多解像度オブジェクトであるかどうかを確認するステップと、もしオブジェクト 'y' が多解像度オブジェクトでなければ、'y' のCDEを除去して 'Sy' バイトから構成されるオブジェクト 'y' へ割り振られたキャッシュ空間の量を解放し、'S_FREE_NEW' = 'S_FREE' + 'Sy' を形成するステップと、'S_FREE_NEW' が 'Sx' よりも大きいかどうかを決定し、もし大きければ、'x' を記憶し、もし大きくなければ、次の使用近時性最小オブジェクト 'y' のために前記確認ステップを反復するステップとを含む、上記（17）に記載のオブジェクトの変更方法。

（19）オブジェクト 'y' が多解像度オブジェクトであるとき、オブジェクト 'y' が前記キャッシュ内の解像度レベル 0' 版で存在するかどうかを決定し、もしそれが存在すれば、前記解像度レベル 0' 位置フィールドをヌルにして、'S_FREE_NEW' = 'S_FREE' + 'Sy' を形成するステップと、'S_FREE_NEW' が 'Sx' よりも大きいかどうかを決定し、もし大きければ、'x' を記憶し、もし大きくなければ、次の使用近時性最小オブジェクト 'y' のために前記確認ステップを反復するステップとを含む、上記（18）に記載のオブジェクトの変更方法。

（20）オブジェクト 'y' が特定の多解像度オブジェクトであるとき、オブジェクト 'y' の解像度レベル 1' 版が前記キャッシュに存在するかどうかを決定し、もしそれが存在すれば、前記解像度レベル 0' 版のためにレベル 1 空間を解放し、前記解像度レベル 0' 版のためにオブジェクト 'y' の前記CDEをヌルにし、'S_FREE' を 'S_FREE' + 'Sy' に増加し、オブジェクト 'y' のCDEを使用近時性最大オブジェクトに設定し、'S_FREE' が 'Sx' よりも大きいかどうかを決定し、もし大きければ、'x' を記憶し、もし大きくなければ、次の使用近時性最小オブジェクト 'y' のために前記確認ステップを反復するステップを含む、上記（18）に記載のオブジェクトの変更方法。

（21）解像度 1 版が発見されなかったとき、オブジェクト 'y' に関連づけられた 'タイプ' に関連づけられた解像度低減プロシージャを、オブジェクト 'y' の前記解像度レベル 0' 版へ適用して、オブジェクト 'y' の解像度

レベル 1' 版を発生するステップと、前記キャッシュ内の 'y' _RES_0 に等しい空間の量を解放し、前記解像度レベル 0' 位置をヌルにするステップと、オブジェクト 'y' を解像度レベル版 'y' _RES_1 で前記キャッシュに記憶するステップと、前記CDE内の解像度 1 位置フィールドが前記レベル 0 位置を指すように設定するステップと、'S_FREE' を 'S_FREE' + 'Sy' _SIZE_0 へ増加するステップと、'S_FREE' を 'S_FREE' - 'Sy' _SIZE_1 に減少するステップと、オブジェクト 'y' の 'y' _CDE を前記使用近時性最大に設定するステップとを含む、上記（20）に記載のオブジェクトの変更方法。

（22）コンピュータ使用可能媒体を含み、複数のオブジェクトをキャッシュに記憶するシステムでキャッシュ空間の量を節約するコンピュータ読み取り可能プログラム・コード手段が前記コンピュータ使用可能媒体に実現された製造物であって、前記コンピュータ読み取り可能プログラム・コード手段は、少なくとも 1 つの前記オブジェクトの少なくとも 1 つの小サイズ低解像度版を発生するステップと、前記低解像度版を前記キャッシュに保存するステップと、前記少なくとも 1 つの前記オブジェクトの大サイズ高解像度版を廃棄するステップとをコンピュータに実行させる、コンピュータ使用可能媒体を含む製造物。

（23）前記オブジェクトの各々が第 1 解像度のデータを有するとき、前記オブジェクトの 1 つの前記第 1 解像度のデータが第 1 解像度削除方針に従っていつ削除されるかを決定するステップを前記コンピュータに実行させる、上記（22）に記載のコンピュータ使用可能媒体を含む製造物。

（24）前記オブジェクトの各々が低解像度のデータを有するとき、前記オブジェクトの 1 つの前記低解像度のデータが低解像度削除方針に従っていつ削除されるかを決定するアクションを前記コンピュータに実行させる、上記（22）に記載のコンピュータ使用可能媒体を含む製造物。

（25）複数のオブジェクトをキャッシュに記憶するシステムで前記有効なキャッシュ空間を節約する方法のステップを実行するために前記機械によって実行可能なプログラム命令を有形的に実現した、機械読み取り可能なプログラム記憶装置であって、前記方法のステップは、少なくとも 1 つの前記オブジェクトの少なくとも 1 つの小サイズ低解像度版を発生するステップと、前記低解像度版をキャッシュに保存するステップと、前記少なくとも 1 つの前記オブジェクトの大サイズ高解像度版を廃棄するステップとを含む、プログラム記憶装置。

（26）前記オブジェクトの各々が第 1 のキャッシュ量を占める第 1 解像度のデータを有するとき、前記方法が、キャッシュの自由空間の量を増加して前記キャッシュに存在しない他のオブジェクトのための余地を作るため、前記オブジェクトの 1 つの前記第 1 解像度のデータ

が第 1 解像度削除方針に従っていつ削除されるかを決定するステップを含む、上記 (25) に記載のプログラム記憶装置。

(27) 前記オブジェクトの各々が小サイズ低解像度のデータを有するとき、前記方法が、オブジェクトの 1 つの前記小サイズ低解像度のデータが低解像度削除方針に従っていつ削除されるかを決定するステップを含む、上記 (25) に記載のプログラム記憶装置。

(28) 各々の多解像度オブジェクトが最初に複数の解像度版で記憶されるとき、現在前記キャッシュに存在しない新しいオブジェクト x を記憶する空間を作成するために使用され、前記順序リストの前記終わりにある最低優先順位オブジェクトに対応する特定のキャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、 y が多解像度オブジェクトであるとき、オブジェクト y の複数の解像度版が現在記憶されているかどうかを確認するステップと、オブジェクト y の複数の解像度版が現在記憶されているとき、自由空間を形成するため前記オブジェクト y の現在の最高解像度版を除去するステップと、オブジェクト y の前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって、オブジェクト y の前記キャッシュ・ディレクトリ・エントリーを最高優先順位オブジェクトに設定するステップと、前記自由空間の量が前記オブジェクト x を記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記自由空間の量がオブジェクト x を記憶するのに十分となるまで、順序リスト内の次の最低優先順位オブジェクトのために決定、確認、除去、設定、及び計算の前記ステップを反復し、オブジェクト x を前記自由空間に記憶するステップとを含む、上記 (2) に記載のオブジェクトの管理方法。

(29) オブジェクト y の 1 つの解像度版のみが現在記憶されているとき、オブジェクト y の低減解像度版を発生及び記憶するステップと、前記自由空間の量を増加するためオブジェクト y の前記現在の高解像度版を除去するステップと、オブジェクト y の前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリーを前記最高優先順位オブジェクトに設定するステップと、前記新しく増加された自由空間が前記オブジェクト x を記憶するのに十分であるかどうかを計算するステップと、もし十分でなければ、前記自由空間の量がオブジェクト x を記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リストにおける前記次の最低優先順位オブジェクトのために決定、確認、発生、除去、設定、配置、及び計算の前記ステップを反復し、オブジェクト x を前記自由空間に記憶するステップと、 y が多解像度オブジェクトでないとき、前記オブジェクト y を除去し、前記新しく増加された自由

空間の量が前記オブジェクト x を記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記順序リスト内の前記次の最低優先順位使用オブジェクトのために決定、確認、発生、除去、及び計算の前記ステップを反復するステップとを含む、上記 (28) に記載のオブジェクトの管理方法。

(30) 複数のオブジェクトをキャッシュに記憶するシステムでキャッシュ空間の量を節約する方法であって、少なくとも 1 つの前記オブジェクトの少なくとも 1 つの小メモリ・サイズ版であってキャッシュ空間の第 1 の量を占める版を発生するステップと、前記小メモリ・サイズ版をキャッシュに保存するステップと、前記少なくとも 1 つの前記オブジェクトの高メモリ・サイズ版であってキャッシュ空間の前記第 1 の量よりも大きいキャッシュ空間の第 2 の量を占める版を廃棄するステップとを含む、キャッシュ空間の節約方法。

(31) キャッシュ内の複数のオブジェクトを管理する方法であって、存在する非多解像度オブジェクト及び少なくとも 1 つの多解像度オブジェクトのためにエントリーを有するキャッシュ・ディレクトリを作成し、各キャッシュ・ディレクトリ・エントリーが前記少なくとも 1 つの多解像度オブジェクトのすべての存在する解像度版を指すポインタを有するようにするステップと、前記エントリーの順序リストを形成するステップとを含む、オブジェクトの管理方法。

(32) 前記順序リストが配列基準に従って形成される、上記 (31) に記載のオブジェクトの管理方法。

(33) 前記配列基準が、特定のオブジェクトが使用された最後の使用時点に基づく、上記 (32) に記載のオブジェクトの管理方法。

(34) 各々の多解像度オブジェクトが最初に複数の解像度版で記憶され、各々の解像度版が異なった量のキャッシュ空間を使用しているとき、現在前記キャッシュに記憶されていない新しいオブジェクト x を記憶する空間を作成するために使用され、前記順序リストの前記終わりにある使用近時性最小オブジェクト y に対応する特定のキャッシュ・ディレクトリ・エントリーが多解像度オブジェクトであるかどうかを決定するステップと、 y が多解像度オブジェクトであるとき、前記オブジェクトの複数の解像度版が現在記憶されているかどうかを確認するステップと、複数の解像度版が現在記憶されているとき、前記オブジェクト y の前記現在の最高解像度版を除去するステップと、オブジェクト y の前記キャッシュ・ディレクトリ・エントリーを前記順序リストの前記始めに配置することによって前記キャッシュ・ディレクトリ・エントリーを前記使用近時性最大オブジェクトに設定するステップと、新しく増加した自由空間の量が前記オブジェクト x を記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記自由空間の量がオブジェクト x を記憶するのに十分となるまで、前記順序リスト

内の前記次の使用近時性最小オブジェクトのために決定、確認、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、上記（３３）に記載のオブジェクトの管理方法。

（３５）１つの解像度版のみが現在記憶されているとき、オブジェクトyの低減解像度版を発生及び記憶するステップと、前記オブジェクトyの前記現在の高解像度版を除去するステップと、前記順序リストの前記始めにオブジェクトyの前記キャッシュ・ディレクトリ・エントリーを配置することによって前記キャッシュ・ディレクトリ・エントリーを前記使用近時性最大オブジェクトに設定するステップと、前記新しく増加した自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを計算するステップと、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、キャッシュ・ディレクトリ・エントリーの前記順序リストにある前記次の使用近時性最小オブジェクトのために決定、確認、発生、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップと、もしyが多解像度版オブジェクトでなければ、前記オブジェクトyを除去し、前記新しく増加した自由空間が前記オブジェクトxを記憶するのに十分であるかどうかを計算し、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、前記順序リスト内の前記次の使用近時性最小オブジェクトのために決定、確認、発生、除去、設定、及び計算の前記ステップを反復し、オブジェクトxを前記自由空間に記憶するステップとを含む、上記（３４）に記載のオブジェクトの管理方法。

（３６）キャッシュ及びキャッシュ・マネージャを備えた装置であって、前記キャッシュ・マネージャは、複数のオブジェクトが前記キャッシュに記憶されるときにキャッシュ空間の量を節約するキャッシュ空間節約手段を含む、装置。

（３７）CPU、ネットワーク・オブジェクト縮減手段、及びメモリ・ディスクを備えた、上記（３６）に記載の装置。

（３８）前記キャッシュがキャッシュ・ディレクトリを備えた、上記（３６）に記載の装置。

（３９）オブジェクトの各々へ割り当てられた優先順位に従って前記オブジェクトの順序リストを形成する前記ステップを含む、上記（１）に記載の記憶域空間の節約方法。

（４０）割り当てられた前記優先順位がオブジェクト使用の頻度に基づく、上記（３９）に記載の記憶域空間の節約方法。

（４１）各オブジェクトを現在の解像度版で記憶するステップと、最低優先順位オブジェクトyが多解像度オブジェクトであるかどうかを確認するステップと、オブジ

ェクトyが多解像度オブジェクトであるとき、オブジェクトyの低減解像度版を発生及び記憶するステップと、自由空間の量を前記メモリに作成するためオブジェクトyの前記現在の解像度版を除去するステップと、オブジェクトyの優先順位を最高優先順位に設定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分でないとき、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、前記メモリ内の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクトxを前記自由空間に記憶するステップとを含む、上記（１）に記載の記憶域空間の節約方法。

（４２）もしyが多解像度オブジェクトでなければ、前記自由空間の量を増加するためにオブジェクトyを前記メモリから除去するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、もし前記自由空間の量がオブジェクトxを記憶するのに十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで前記メモリ内の次の最低優先順位オブジェクトのために確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクトxを前記自由空間に記憶するステップとを含む、上記（４１）に記載の記憶域空間の節約方法。

（４３）確認する前記ステップが前記オブジェクトの順序リストのメモリ・ディレクトリ・エントリーを形成するステップを含む、上記（４１）に記載の記憶域空間の節約方法。（４４）使用近時性最小オブジェクトが多解像度オブジェクトであるかどうかを確認するステップと、もしオブジェクトyが多解像度オブジェクトであれば、オブジェクトyの低減解像度版を発生及び記憶するステップと、自由空間を作成するように、オブジェクトyの現在の高解像度版を前記メモリから除去するステップと、オブジェクトyの優先順位を使用近時性最大オブジェクトに設定するステップと、前記自由空間の量がオブジェクトxを記憶するのに十分であるかどうかを決定するステップと、もし十分でなければ、前記自由空間の量がオブジェクトxを記憶するのに十分となるまで、確認、発生、除去、設定、及び決定の前記ステップを反復するステップと、オブジェクトxを前記自由空間に記憶するステップとを含む、上記（１）に記載の記憶域空間の節約方法。

（４５）前記メモリが一時的バッファである、上記（１）に記載の記憶域空間の節約方法。

（４６）前記メモリがキャッシュである、上記（１）に記載の記憶域空間の節約方法。

（４７）コンピュータが多解像度オブジェクトに対する要求を満足させる方法であって、多解像度オブジェクト

10

20

30

40

50

を、要求された解像度レベルで渡すべき旨の前記要求を受け取るステップと、前記コンピュータが、前記要求された解像度レベルの前記オブジェクトを有するかどうかを決定するステップと、前記コンピュータが、もし前記要求された解像度レベルの前記オブジェクトを有するならば、前記要求されたオブジェクトを渡すステップと、前記コンピュータが、もし前記要求された解像度よりも高い解像度レベルの前記要求されたオブジェクトを有するならば、前記要求された解像度レベルの前記オブジェクトを発生するステップと、前記オブジェクトを渡すステップとを含む、オブジェクトに対する要求を満足させる方法。

【図面の簡単な説明】

【図 1】本発明がクライアント・コンピュータ、プロキシ・コンピュータ、及びサーバ・コンピュータの各々で実現される場合に、これらコンピュータの相互接続シナリオを示す図である。

【図 2】図 1 に示されるサーバ・コンピュータ、プロキシ・コンピュータ、及びクライアント・コンピュータの内部構成要素を示すブロック図である。

【図 3】ネットワーク通信スタックの層、及びそれらの層とキャッシュ・マネージャの相互作用を示すブロック図である。

【図 4】典型的にネットワーク・オブジェクトのキャッシュ・ディレクトリ・エントリを有するキャッシュ・ディレクトリを使用し、オブジェクトが記憶されているキャッシュ・メモリ位置を指すポインタをキャッシュ・ディレクトリ・エントリが含んでいるネットワーク・オブジェクト・キャッシュを示す図である。

【図 5】図 4 のキャッシュ・ディレクトリにあるエントリの詳細図である。

【図 6】新しいネットワーク・オブジェクトを記憶する空間を作成するためにネットワーク・キャッシュからオブジェクトを除去する方法を示すフローチャートである。

【図 7】本発明に従った多解像度ネットワーク・オブジェクトの場合に、図 4 のキャッシュ・ディレクトリにあるエントリを示す詳細図である。

【図 8】オブジェクトの解像度版が一時に 1 つだけキャッシュに記憶される場合に、本発明に従って新しいネットワーク・オブジェクトを記憶する空間を作成するためにネットワーク・キャッシュ内の多解像度オブジェクトを縮減する方法を示す高レベルの図である。

【図 9】オブジェクトの多解像度版が一時に複数個キャッシュに記憶される場合に、本発明に従って新しいネットワーク・オブジェクトを記憶する空間を作成するためにネットワーク・キャッシュ内の多解像度オブジェクトを縮減する方法を示す高レベルの図である。

【図 10】オブジェクトの版が一時に 1 つだけキャッシュに存在する場合に、本発明に従って新しいネットワー

ク・オブジェクトを記憶する空間を作成するためにネットワーク・キャッシュ内の非多解像度オブジェクトを除去すると共に多解像度オブジェクトを縮減又は除去する方法を示すフローチャートである。

【図 11】オブジェクトの複数の解像度レベルがキャッシュにあらかじめ存在する場合に、本発明に従って新しいネットワーク・オブジェクトを記憶する空間を作成するためにネットワーク・キャッシュ内の非多解像度オブジェクトを除去すると共に多解像度オブジェクトを縮減又は除去する方法を示すフローチャートである。

【図 12】オブジェクトが多解像度タイプであるかどうか、もしそうであれば、どの解像度低減プロシージャを本発明に従って使用するかを決定するテーブルを示す図である。

【符号の説明】

1 0 0	クライアント・コンピュータ
1 6 0	プロキシ
1 7 0	ウェブ・サーバ
2 1 1	C P U
2 1 2	特定ハードウェア構成要素
2 1 4	キャッシュ・ディレクトリ
2 1 5	ネットワーク・オブジェクト・キャッシュ
2 1 6	メイン・メモリ
2 1 7	キャッシュ・メモリ
2 1 8	ネットワーク・オブジェクト解像度低減プロシージャ
2 1 9	アプリケーション・プログラム
2 2 0	キャッシュ・マネージャ
2 5 0	プロキシ・アプリケーション
2 5 1	H T T P プロトコル層
2 5 2	T C P / I P プロトコル群
2 5 4	キャッシュ・マネージャ
2 5 5	キャッシュ・メモリ
3 0 0	キャッシュ・ディレクトリ
3 1 0	キャッシュ・メモリ
3 2 0	キャッシュ・ディレクトリ・エントリ
(C D E)	
3 5 0	ネットワーク・オブジェクト
4 0 0	C D E
4 1 0	フィールド
4 1 5	フィールド
4 2 0	フィールド
4 2 5	フィールド
4 3 0	フィールド
4 4 0	フィールド
4 5 0	フィールド
6 1 0	フィールド
6 2 0	フィールド
6 3 0	ポインタ
6 4 0	ポインタ

10

20

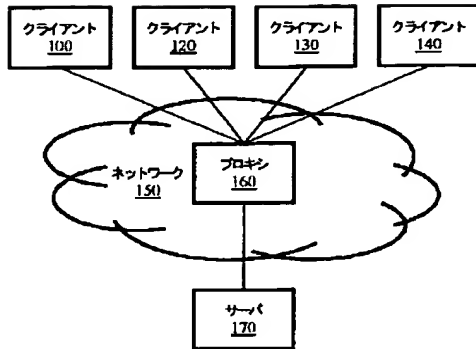
30

40

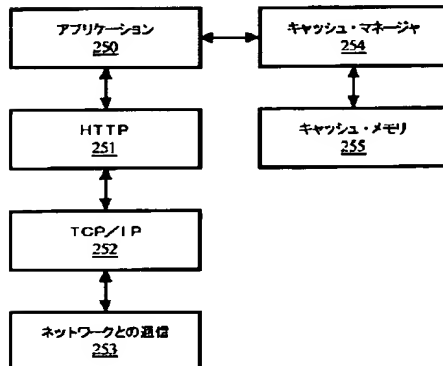
50

6 5 0 フィールド
 6 6 0 フィールド
 6 7 0 フィールド
 6 8 0 C D E
 6 8 3 C D E
 6 8 4 オブジェクト

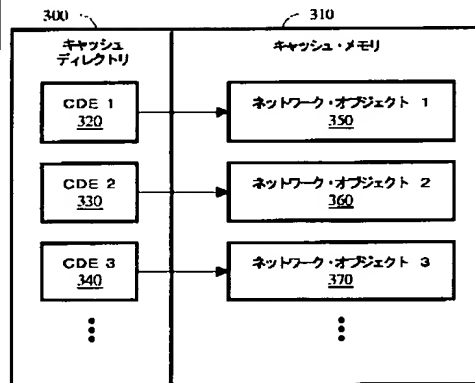
【図 1】



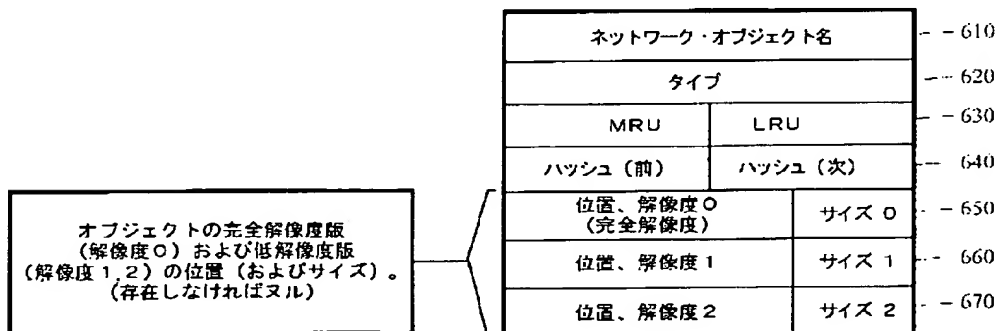
【図 3】



【図 4】

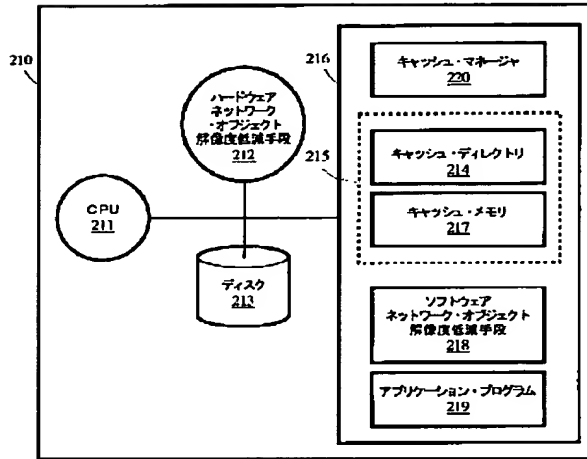


【図 7】



6 8 5 低解像度版
 6 9 0 C D E
 6 9 3 C D E
 6 9 4 オブジェクト
 8 1 0 テーブル・ヘッダ
 8 2 0 エントリー

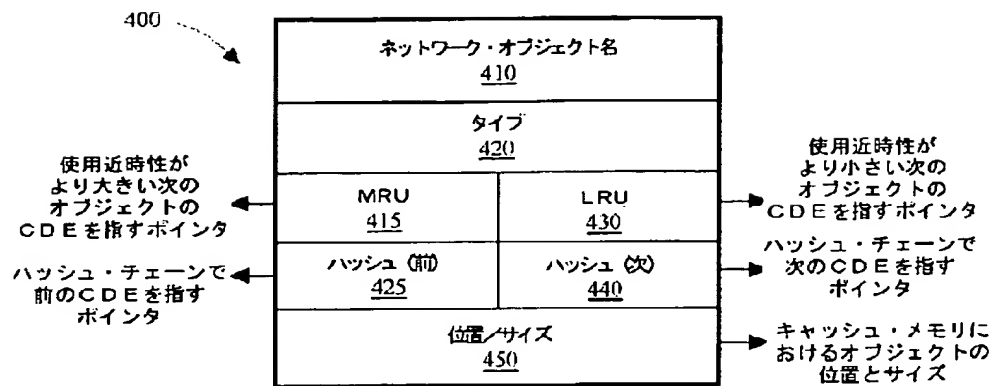
【図 2】



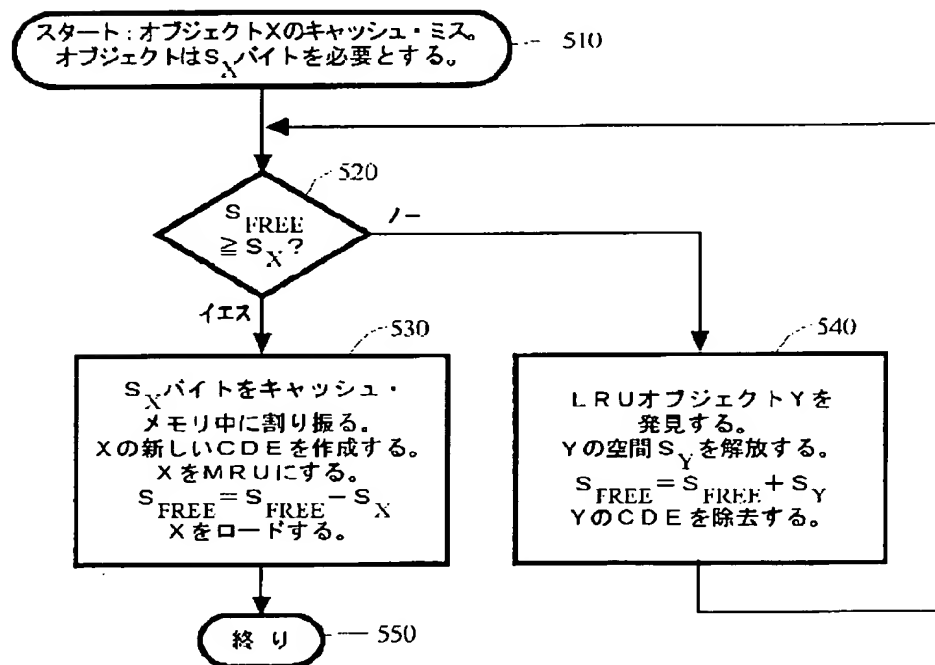
【図 1 2】

タイプの数 (3)		810
タイプ名 1	解像度低減 プロシージャ 1	820
タイプ名 2	解像度低減 プロシージャ 2	830
タイプ名 3	解像度低減 プロシージャ 3	840

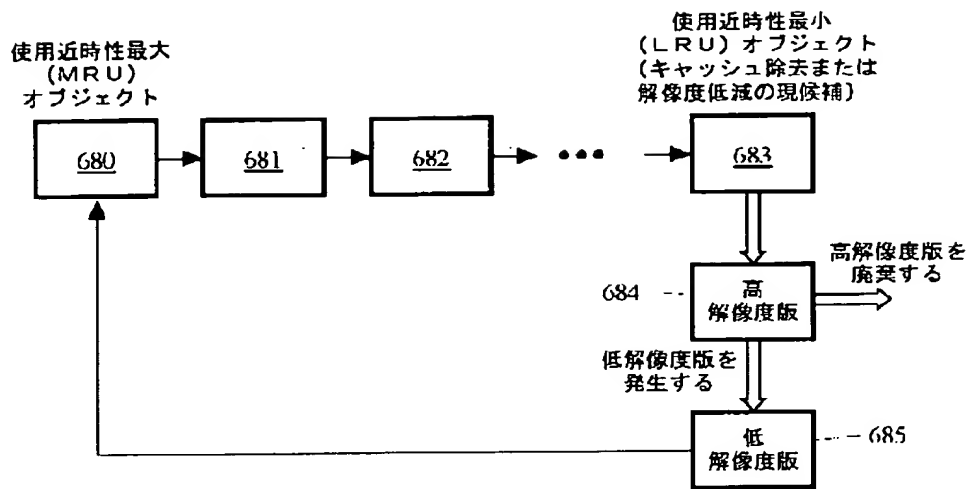
【図 5】



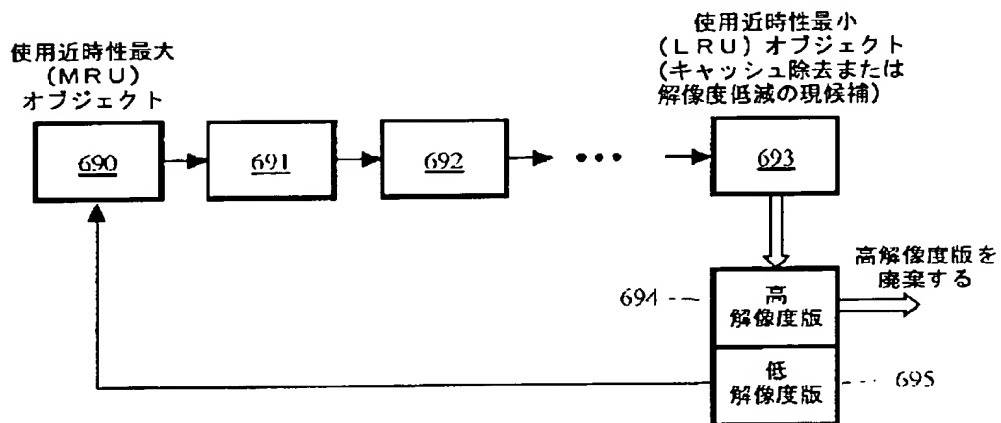
【図 6】



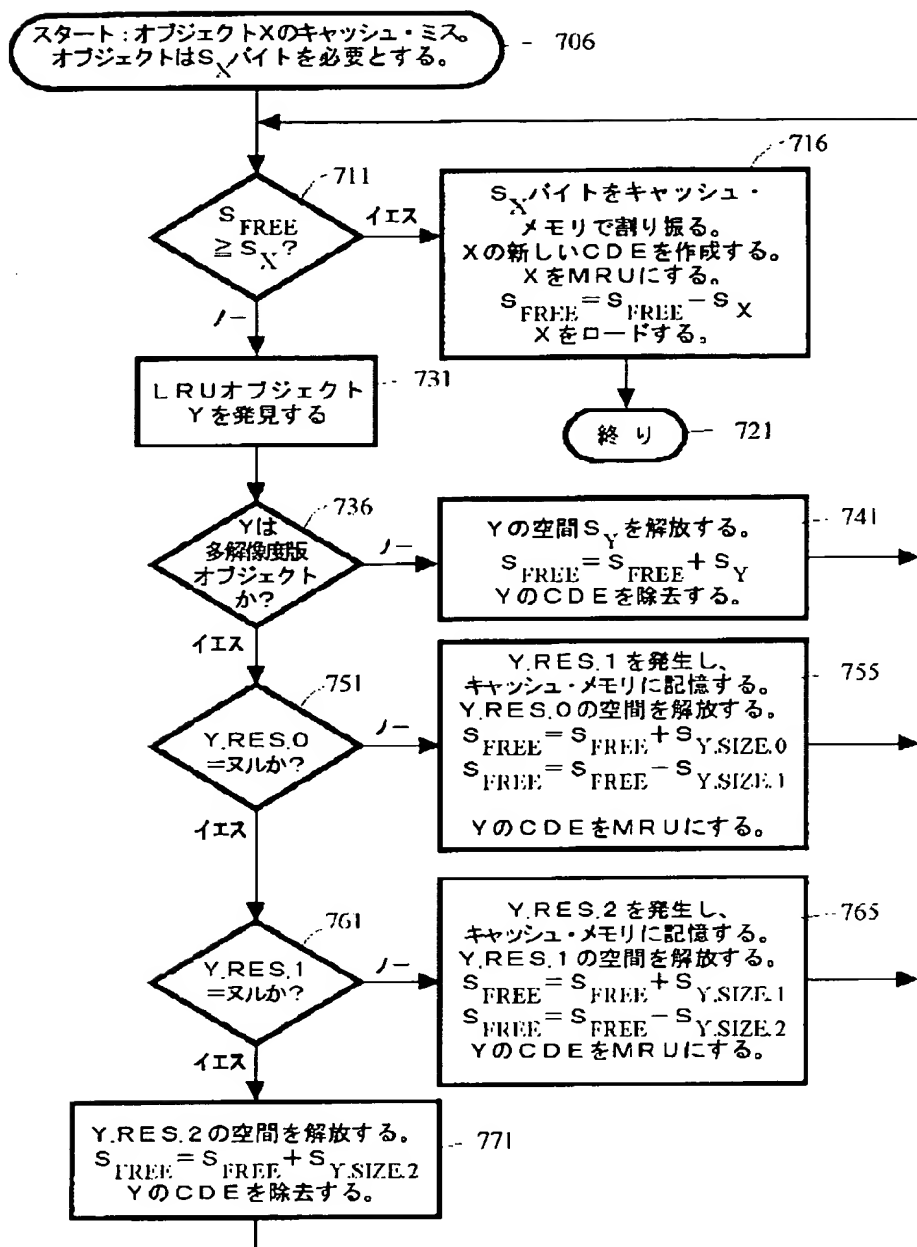
【図 8】



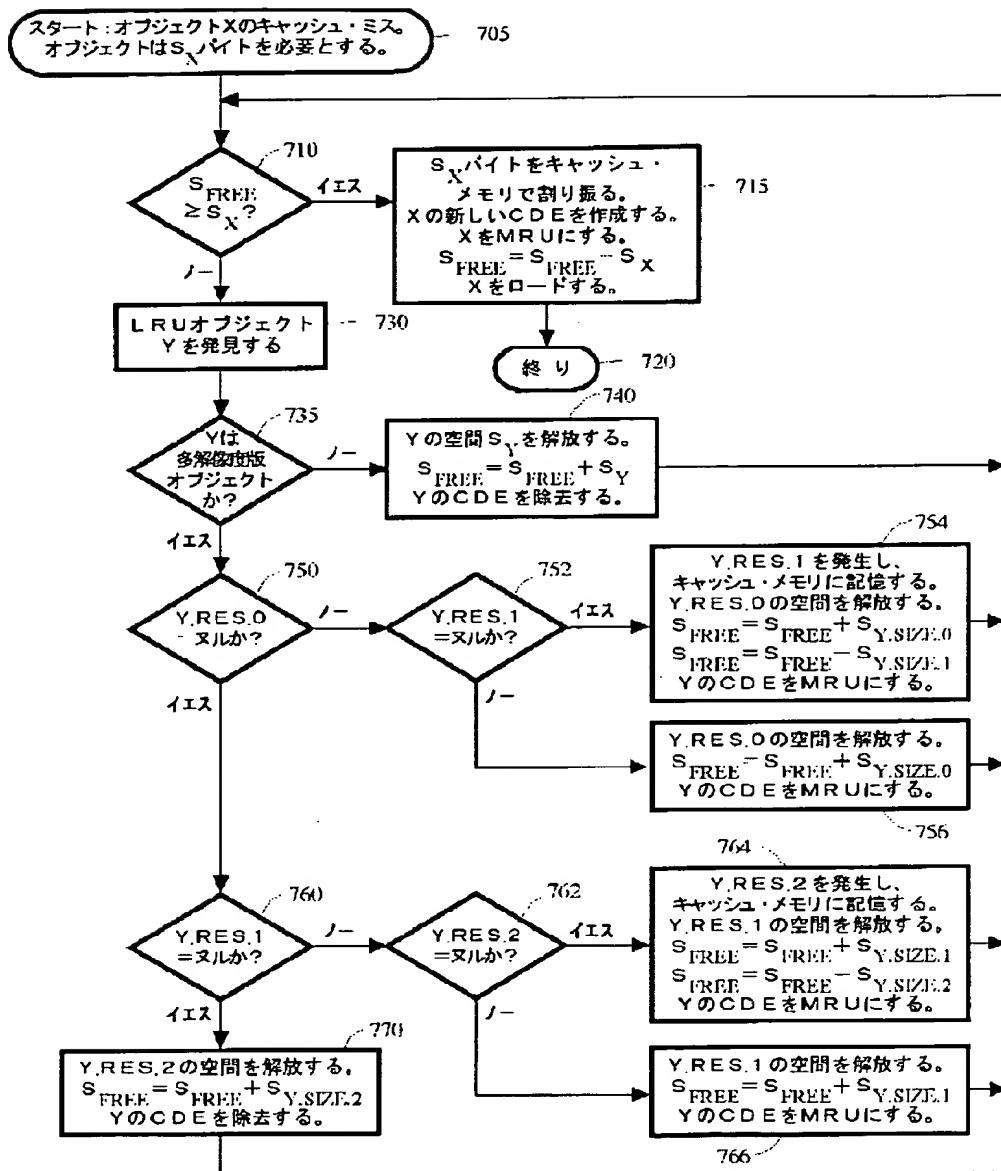
【図 9】



【図 10】



【図 1 1】



フロントページの続き

(72)発明者 ジョン・ティモシー・ロビンソン
 アメリカ合衆国10598 ニューヨーク州、
 ヨークタウン・ハイツ、 ノース・ディ
 アフィールド・アヴェニュー 3314